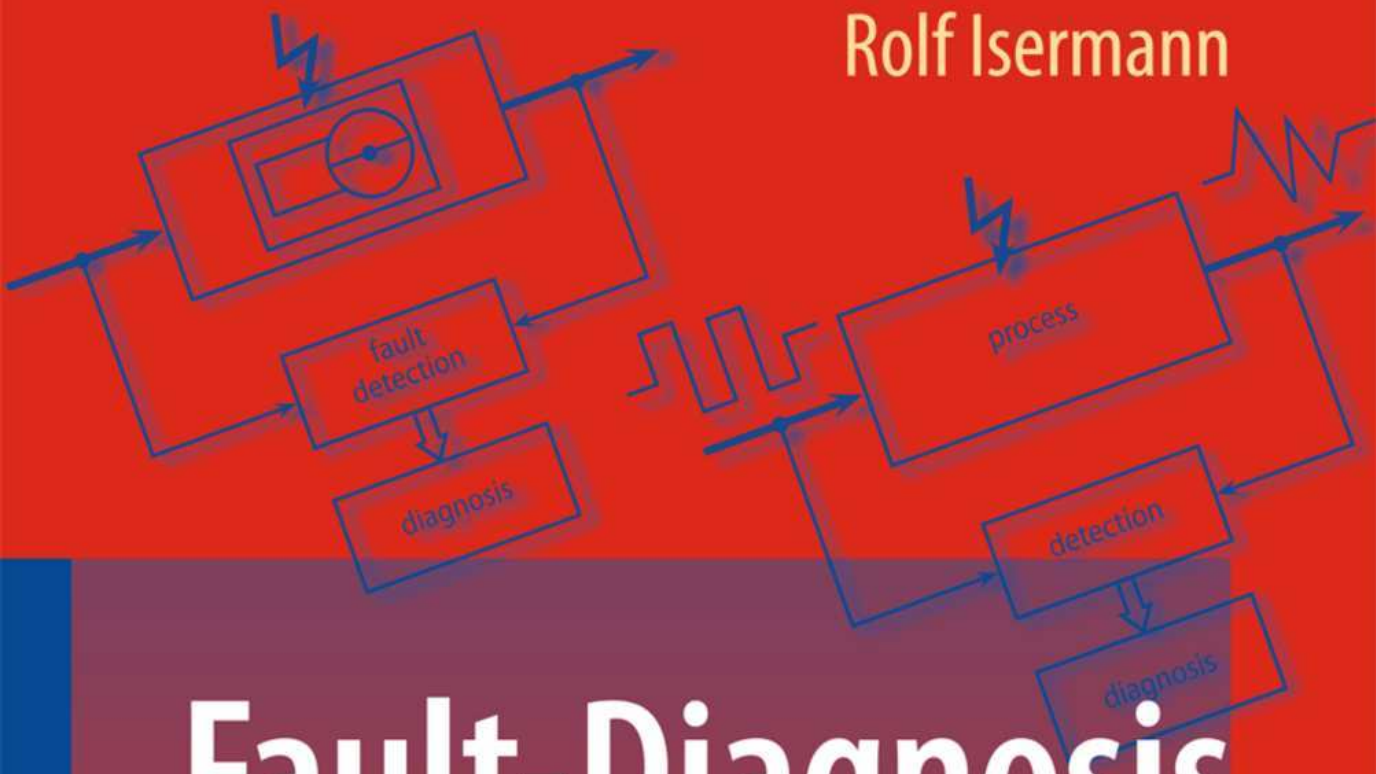
Rolf Isermann



Fault-Diagnosis Applications

Model-based condition monitoring: Actuators, drives, machinery, plants, sensors, and fault-tolerant systems

Springer

VISIT…

Fault-Diagnosis Applications

Rolf Isermann

Fault-Diagnosis Applications

Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems

Springer

Prof. Dr. -Ing. Dr. h.c. Rolf Isermann
Technische Universität Darmstadt
Institut für Automatisierungstechnik
Landgraf-Georg-Straße 4
64283 Darmstadt
Germany
risermann@iat.tu-darmstadt.de

ISBN 978-3-642-12766-3 e-ISBN 978-3-642-12767-0
DOI 10.1007/978-3-642-12767-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011926417

© Springer-Verlag Berlin Heidelberg 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: WMXDesign GmbH

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

With increasing demands for efficiency and product quality and progressing integration of automatic control systems in high-cost and safety-critical processes, the field of supervision (or monitoring), fault detection and fault diagnosis plays an important role. The classical method of supervision is to check the limits of single variables and alarming of operators. However, this can be improved significantly by taking into account the information hidden in all measurements and by automatic actions to keep the systems in operation.

During the last few decades theoretical and experimental research has shown new ways to detect and diagnose faults. One distinguishes *fault detection* to recognize that a fault happened, and *fault diagnosis* to find the cause and location of the fault. Advanced methods of fault detection are based on mathematical signal and process models and on methods of system theory and process modeling to generate fault symptoms. Fault-diagnosis methods use causal fault–symptom relationships by applying methods from statistical decision, artificial intelligence and soft computing. Therefore, efficient supervision, fault detection and diagnosis is a challenging field encompassing physical-oriented system theory, experiments and computations. The considered subjects are also known as condition monitoring, fault detection and isolation (FDI) or fault detection and diagnosis (FDD).

A further important field is *fault management* or *asset management*. This means to avoid shutdowns by early fault detection and actions like *process condition-based maintenance* or *repair*. If sudden faults, failures or malfunctions cannot be avoided, *fault-tolerant systems* are required. Through methods of fault detection and reconfiguration of redundant components, breakdown, and in the case of safety-critical processes, accidents, may be avoided.

As the successor to the book *Fault-Diagnosis Systems – An Introduction from Fault Detection to Fault Tolerance* this book describes *applications of fault detection and diagnosis* to different kinds of technical processes and products.

The development of fault-detection and fault-diagnosis methods was paralleled by experimental investigations with several technical processes at the Institute of Automatic Control of the Darmstadt University of Technology from about 1975 and frequently in cooperation with industrial partners. In this way the theoretically de-

veloped methods could be tested on real processes and the experimental results gave hints for improvements and further ideas. Therefore, this book contains the main results of 20 different processes, ranging from electrical drives and different actuators, through machine tools, pumps and pipelines, to heat exchangers. (Theoretical and practical results for combustion engines are published in another book *Engine Control and Diagnosis*.)

The book is an introduction to the application of fault diagnosis and fault tolerance in the areas of electrical engineering, mechanical and chemical engineering and computer science. It is addressed to students and practicing engineers in research and development, design and manufacturing. Preconditions are basic undergraduate courses in system theory, automatic control, and mechanical and/or electrical engineering.

The author is grateful to his research associates, who have performed many theoretical and practical research projects on the subject of this book since 1975, among them H. Siebert, L. Billmann, G. Geiger, W. Goedecke, S. Nold, U. Raab, B. Freyermuth, St. Leonhardt, R. Deibert, T. Höfling, T. Pfeufer, M. Ayoubi, P. Ballé, D. Füssel, O. Moseler, A. Wolfram, M. Münchhof, F. Haus and M. Beck.

Finally, I especially would like to thank Brigitte Hoppe for the laborious and precise text setting, including the figures and tables in camera-ready form and Springer-Verlag for the excellent cooperation.

Darmstadt, September, 2010

Rolf Isermann

Contents

List of Symbols XIII

1 Introduction 1
1.1 Process automation and process supervision/condition monitoring .. 1
1.2 Product life cycles and fault management (asset management) 3
1.3 Contents 5

Part I Supervision, Fault Detection and Diagnosis

2 Supervision, fault-detection and diagnosis methods – a short introduction 11
2.1 Basic tasks of supervision 11
2.2 Terminology 17
2.2.1 Faults, failures, malfunctions 17
2.2.2 Reliability, availability, safety 19
2.2.3 Fault tolerance and redundancy 21
2.3 Knowledge-based fault detection and diagnosis 22
2.3.1 Analytic symptom generation 22
2.3.2 Heuristic symptom generation 23
2.3.3 Fault diagnosis 24
2.4 Signal-based fault-detection methods 24
2.4.1 Limit checking of absolute values 25
2.4.2 Trend checking 25
2.4.3 Change detection with binary thresholds 26
2.4.4 Adaptive thresholds 27
2.4.5 Plausibility checks 28
2.4.6 Signal-analysis methods 29
2.5 Process-model-based fault-detection methods 30
2.5.1 Process models and fault modeling 32
2.5.2 Fault detection with parameter estimation 34

2.5.3 Fault detection with state observers and state estimation 35
2.5.4 Fault detection with parity equations 37
2.5.5 Direct reconstruction of non-measurable variables 38
2.6 Fault-diagnosis methods 39
2.6.1 Classification methods 39
2.6.2 Inference methods 41
2.7 Fault detection and diagnosis in closed loop 41
2.8 Data flow structure for supervision (condition monitoring) 43

Part II Drives and Actuators

3 Fault diagnosis of electrical drives 49
3.1 Direct-current motor (DC) 49
3.1.1 Structure and models of a DC motor 49
3.1.2 Fault detection with parity equations 52
3.1.3 Fault detection with parameter estimation 54
3.1.4 Experimental results for fault detection (SELECT) 55
3.1.5 Experimental results for fault diagnosis with a learning fault-symptom tree 57
3.1.6 Conclusions 63
3.2 Alternating-current motor (AC) 63
3.2.1 Structure and models of induction motors (asynchronous motors) 64
3.2.2 Signal-based fault detection of the power electronics 66
3.2.3 Model-based fault detection of the AC motor 71
3.2.4 Conclusions 80

4 Fault diagnosis of electrical actuators 81
4.1 Electromagnetic actuator 81
4.1.1 Position control 83
4.1.2 Fault detection with parameter estimation 85
4.2 Electrical automotive throttle valve actuator 87
4.2.1 Structure and models of the actuator 88
4.2.2 Input test cycle for quality control 89
4.2.3 Fault detection with parameter estimation 91
4.2.4 Fault detection with parity equation 94
4.2.5 Fault diagnosis 95
4.2.6 Fault-diagnosis equipment 96
4.2.7 Conclusions 98
4.3 Brushless DC motor and aircraft cabin pressure valve 98
4.3.1 Structure and models 98
4.3.2 Fault detection with parameter estimation 101
4.3.3 Fault detection with parity equations 102
4.3.4 Conclusions 104

5 Fault diagnosis of fluidic actuators 105
5.1 Hydraulic servo axis 105
5.1.1 Hydraulic servo axis structure 106
5.1.2 Faults of hydraulic servo axes 106
5.1.3 Models of spool valve and cylinder 111
5.1.4 Fault detection and diagnosis of valve and cylinder 115
5.1.5 Conclusions 121
5.2 Pneumatic actuators 121
5.2.1 Pneumatic-actuator construction 122
5.2.2 Faults of pneumatic valves 124
5.2.3 Models of pneumatic valves 125
5.2.4 Fault detection with valve characteristics 128
5.2.5 Fault detection of flow valves with pneumatic position controller 130
5.2.6 Fault detection of flow valves with electronic position controller 138
5.2.7 Conclusions 139

Part III Machines and Plants

6 Fault diagnosis of pumps 143
6.1 Centrifugal pumps 143
6.1.1 State of the art in pump supervision and fault detection 143
6.1.2 Models of centrifugal pumps and pipe systems 146
6.1.3 Fault detection with parameter estimation 149
6.1.4 Fault detection with nonlinear parity equations and parameter estimation 156
6.1.5 Fault detection with vibration sensors 165
6.1.6 Conclusions 169
6.2 Reciprocating pumps 170
6.2.1 Structure of a diaphragm pump 171
6.2.2 Models of a diaphragm pump 172
6.2.3 Fault detection and fault diagnosis of the hydraulic pump . . . 172
6.2.4 Fault detection of the pump drive 177
6.2.5 Conclusions 178

7 Leak detection of pipelines 181
7.1 State of the art in pipeline supervision 181
7.2 Models of pipelines 182
7.3 Model-based leak detection 187
7.3.1 Leak detection with state observers 188
7.3.2 Leak detection with mass balance and correlation analysis for liquid pipelines 190
7.3.3 Leak detection for gas pipelines 195

7.4 Experimental results 201
7.4.1 Gasoline pipeline 201
7.4.2 Gas pipeline 202
7.4.3 Conclusions 204

8 Fault diagnosis of industrial robots 205
8.1 Structure of a six-axis robot 205
8.2 Model of a robot axis and parameter estimation 206
8.3 Analytic and heuristic diagnosis knowledge 208
8.3.1 Symptom representation 208
8.3.2 Diagnosis knowledge representation 209
8.3.3 Faults, heuristic symptoms and events of the robot 210
8.4 Experimental results 211
8.4.1 Fault diagnosis with analytical knowledge 212
8.4.2 Fault diagnosis with analytical and heuristic knowledge 213
8.5 Conclusions 215

9 Fault diagnosis of machine tools 217
9.1 Structures of machine tools 217
9.2 Status of machine tools supervision 219
9.3 Main drive 221
9.3.1 Two-mass model 221
9.3.2 Parameter estimation 223
9.3.3 Fault detection with parameter estimation 225
9.4 Feed drives 226
9.4.1 Two- and three-mass model 226
9.4.2 Identification of a feed drive 229
9.4.3 Fault detection of a feed drive test rig 229
9.5 Drilling machines 234
9.5.1 Models of the drilling process 234
9.5.2 Fault detection of drilling 237
9.6 Milling machine 239
9.6.1 Models for the milling process 239
9.6.2 Fault detection of the cutter 245
9.7 Grinding machines 251
9.7.1 Grinding-process models 252
9.7.2 Fault detection with parameter estimation 254
9.7.3 Fault detection with signal-analysis methods 256
9.8 Conclusions 256

10 Fault detection of heat exchangers 259
10.1 Heat exchangers and their models 259
10.1.1 Heat exchanger types 259
10.1.2 Heat exchanger models for stationary behavior 262
10.1.3 Dynamic models of heated tubes 264

10.2 Fault detection for static behavior 270
10.2.1 Static models of heat exchangers 270
10.2.2 Fault-detection methods 271
10.3 Fault detection for a steam/water heat exchanger with dynamic models and parameter estimation 274
10.3.1 Fault detection with linear dynamic models and parameter estimation 274
10.3.2 Fault detection with parameter variable local linear dynamic models 278
10.4 Conclusions 281

Part IV Fault-tolerant Systems

11 Fault-tolerant systems – a short introduction 285
11.1 Basic redundant structures 285
11.2 Degradation steps 287

12 Examples of fault-tolerant systems 291
12.1 A fault-tolerant control system 291
12.2 Fault-tolerant electrical drives 294
12.2.1 A fault-tolerant duplex AC motor 294
12.2.2 Fault-tolerant frequency converter 298
12.2.3 Multi-phase motors 301
12.3 Fault-tolerant actuators 301
12.3.1 Fault-tolerant hydraulic actuators 302
12.3.2 Fault-tolerant DC actuator 306
12.4 Fault-tolerant sensors 308
12.4.1 Hardware sensor redundancy 308
12.4.2 Analytical sensor redundancy 308
12.4.3 Steering angle sensor 311
12.4.4 Fault-tolerant flow sensor 313
12.4.5 Electronic throttle 314
12.4.6 Virtual drive dynamic sensors by model based analytical redundancy 314

Part V Appendix

13 Terminology in fault detection and diagnosis 321

Concluding remarks 325

References 329

Index 351

List of Symbols

Only frequently used symbols and abbreviations are given.

Letter symbols

a	parameters of differential of difference equations
b	parameters of differential or difference equations
c	spring constant, constant, concentration, stiffness
d	damping coefficient
e	equation error, control deviation $e = w - y$, number $e = 2.71828\ldots$
f	fault, frequency ($f = 1/T_p$, T_p period), function $f(\ldots)$
g	gravitational acceleration, function $g(\ldots)$, impulse response
i	integer, gear ratio, index, $\sqrt{-1}$ (imaginary unit)
j	integer, index
k	discrete number, discrete time $k = t/T_0 = 0, 1, 2, \ldots$ (T_0: sampling time)
l	index, length
m	mass, order number
n	rotational speed, order number, disturbance signal
p	pressure, index, controller parameter, probability density function, process parameter
q	controller parameter, failure density
r	index, radius, reference variable, residual
s	Laplace variable $s = \delta + i\omega$, symptom
t	continuous time, principal component coordinate
u	input signal change ΔU
v	speed, specific volume, disturbance signal
w	reference value, setpoint
x	space coordinate, state variable
y	output signal change ΔY, space coordinate, control variable change ΔY, signal

z	space coordinate, disturbance variable change ΔZ, z-transform: variable $z = \exp T_0 s$
$\hat{x}$	estimated or observed variable
$\tilde{x}$	estimation error
$\bar{x}$	average, steady-state value
x_0	amplitude
x_{00}	value in steady state
A	area
B	magnetic flux density
C	capacitance
D	damping ratio, diameter
E	module of elasticity, energy, potential, bulk modulus
F	filter transfer function, force
G	weight, transfer function
H	magnetic field strength, height
I	electrical current, mechanical momentum, torsion, second moment of area
J	moment of inertia
K	constant, gain
L	inductance
M	torque
N	discrete number, windings number
P	power, probability
Q	generalized force, heat, unreliability function
R	electrical resistance, covariance or correlation function, reliability function, risk number
S	spectral density, sum, performance criterion
T	absolute temperature, torque, time constant
T_0	sampling time
U	input variable, manipulated variable (control input), voltage
V	volume
X	space coordinate
Y	output variable, space coordinate, control variable
Z	space coordinate, disturbance variable
$\mathbf{a}$	vector
$\mathbf{A}$	matrix
$\mathbf{A}^T$	transposed matrix
$\mathbf{I}$	identity matrix
$\boldsymbol{\theta}$	parameter vector
$\mathbf{P}$	covariance matrix
$\boldsymbol{\psi}$	data vector
α	coefficient, angle
β	coefficient, angle

γ	specific weight, correcting factor
δ	decay factor, impulse function
ϕ	correlation function
η	efficiency
ϑ	temperature
λ	thermal conductivity, forgetting factor, failure rate
μ	friction coefficient, permeability, membership function
ν	kinematic viscosity, index
π	number $\pi = 3.14159\ldots$
ρ	density
σ	standard deviation, σ^2 variance
τ	time
φ	angle
ω	angular frequency, $\omega = 2\pi/T_p$; T_p period
Δ	change, deviation
θ	parameter
Π	product
Σ	sum
Θ	magnetic flux
Ψ	magnetic flux linkage

Mathematical abbreviations

$\exp(x) = e^x$	
$E\{\}$	expectation of a statistical variable
dim	dimension
adj	adjoint
det	determinant
Re	real part
Im	imaginary part
$\dot{Y}$	dY/dt (first derivative)
var [ ]	variance
cov [ ]	covariance
$\mathcal{F}$	Fourier transform

Abbreviations

ACF	Auto Correlation Function
ARMA	Auto Regressive Moving Average process
CCF	Cross Correlation Function
DFT	Discrete Fourier Transform
ETA	Event Tree Analysis
FDD	Fault Detection and Diagnosis
FDI	Fault Detection and Isolation

FMEA	Failure Mode and Effects Analysis
FFT	Fast Fourier Transform
FTA	Fault Tree Analysis
HA	Hazard Analysis
IGBT	Insulated Gate Bipolar Transistor
LS	Least Squares
MLP	Multilayer Perceptron
MTBF	Mean Time Between Failures
MTTF	Mean Time To Failure $= 1/\lambda$
MTTR	Mean Time To Repair
NN	Neural Net
PCA	Principal Component Analysis
PRBS	Pseudo Random Binary Signal
RBF	Radial Basis Function
RLS	Recursive Least Squares
rms (...)	root of mean squared of (...)

Part I

Supervision, Fault Detection and Diagnosis

2

Supervision, fault-detection and diagnosis methods – a short introduction [1]

The *supervision* of technical processes and the *quality control* of products is aimed at showing the present state (*condition monitoring*), indicating undesired or unpermitted states, and taking appropriate actions to avoid damage or accidents. The deviations from normal process behavior result from *faults* and *errors*, which can be attributed to many causes. They may result sooner or later in malfunctions or failures if no counteractions are taken. One reason for supervision and quality control is to avoid these malfunctions or failures.

The basic tasks of supervision, fault detection and fault management were already described in detail in a previous volume, [2.37]. Therefore, only some important issues and some basic methods of fault detection and diagnosis which will be used in this book are repeated in this chapter. The condensed presentation of the methods is limited to linear processes. However, these methods can be expanded to nonlinear processes as shown in the previous volume and in many of the following application examples.

2.1 Basic tasks of supervision

A process or a product P which operates in open loop is considered, Figure 2.1a). $U(t)$ and $Y(t)$ are input and output signals, respectively. A fault can now appear due to external or internal causes. Examples for external causes are environmental influences like humidity, dust, chemicals, electromagnetic radiation, high temperature, leading, e.g. to corrosion or pollution. Examples for internal causes are missing lubrication and therefore higher friction or wear, overheating, leaks, and shortcuts. These faults $\mathbf{F}(t)$ firstly affect internal process parameters $\boldsymbol{\Theta}$ by $\Delta\boldsymbol{\Theta}(t)$ like changes of resistance, capacitance or stiffness and/or internal state variables $\mathbf{x}(t)$ by $\Delta\mathbf{x}(t)$ like changes of mass flows, currents or temperatures, which are frequently not measurable. According to the dynamic process transfer behavior, these faults $F(t)$ influence

[1] This chapter is a shortened version of Chapter 2 in [2.37]

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_2, © Springer-Verlag Berlin Heidelberg 2011

the measurable output $Y(t)$ by a change $\Delta Y(t)$. However, it has to be taken into account that also natural process disturbances and noise $N(t)$ and also changes of the manipulated variable $U(t)$ influence the output $Y(t)$.

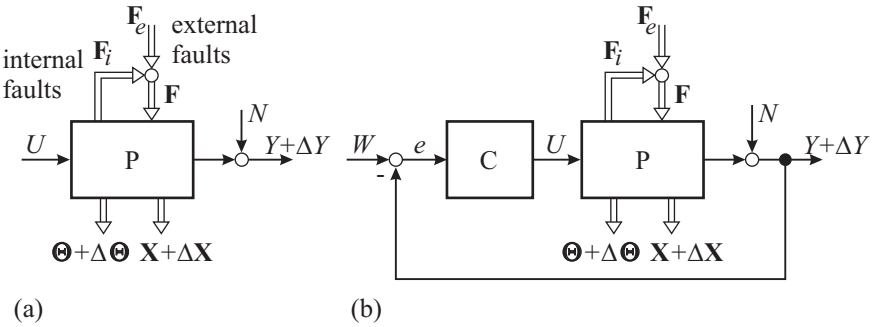


Fig. 2.1. Scheme of a process or product P influenced by faults $\mathbf{F}$: a) process in open loop; b) process in closed loop

For a process operating in open loop a remaining fault $f(t)$ generally results in a permanent offset of $\Delta Y(t)$, as shown in Figure 2.2a). In the case of a *closed loop*, 2.1b), the behavior is different, Figure 2.2b). Depending on the time history of parameter changes $\Delta\boldsymbol{\Theta}(t)$ or state-variable changes $\Delta\mathbf{x}(t)$ the output shows only a somewhat shorter and vanishing small deviation $\Delta Y(t)$ if a controller with integral behavior (e.g. a PI-controller) is used. But then the manipulated variable shows a permanent offset $\Delta U(t)$ for proportionally acting processes. If only the output $Y(t)$ is supervised, the fault may not be detected because of the small and short deviation, furthermore corrupted by noise. The reason is that a closed loop is not only able to compensate for disturbances $N(t)$ but also to compensate for parameter changes $\Delta\boldsymbol{\Theta}(t)$ and state changes $\Delta\mathbf{x}(t)$ with regard to the control variable $Y(t)$. This means that faults $\mathbf{F}(t)$ may be compensated by the closed loop. Only if the fault grows in size and causes the manipulated variable to reach a restriction value (saturation) may a permanent deviation ΔY arise. Hence, for processes in closed loop $U(t)$ should be monitored, as well as $Y(t)$, which is frequently not realized. Mostly, only $Y(t)$ and the control deviation $e(t)$ are supervised.

The *supervision* of technical processes in normal operation or the *quality control* of products in manufacturing is usually performed by *limit checking* or *threshold checking* of some few measurable output variables $Y(t)$, like pressures, forces, liquid levels, temperatures, speeds, and oscillations. This means one checks if the quantities are within a tolerance zone $Y_{min} < Y(t) < Y_{max}$. If the tolerance zone is exceeded, an alarm is raised. Hence, the first task in supervision is, Figure 2.3:

1. *Monitoring*: Measurable variables are checked with regard to tolerances, and alarms are generated for the operator. After an alarm is triggered the operator then

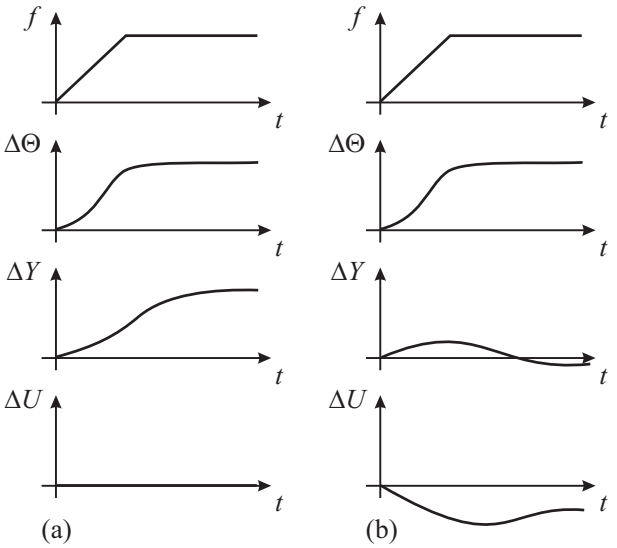


Fig. 2.2. Time behavior of a parameter change $\Delta\Theta$ and measurable signals $Y(t)$ and $U(t)$ after appearance of fault f: a) open loop; b) closed loop

has to take appropriate counteractions.

However, if exceeding a threshold implies a dangerous process state, the counteraction should be generated automatically. This is a second task of supervision, Figure 2.3:

2. *Automatic protection*: In the case of a dangerous process state, the monitoring function automatically initiates an appropriate counteraction. Usually, the process is then commanded to a fail-safe state, which is frequently an emergency shutdown. Some examples are shown in [2.37], Chapter 2.

These *classical methods* of monitoring and automatic protection are suitable for the overall supervision of the processes. To set the tolerances, compromises have to be made between the detection size for abnormal deviations and unmeasurable or wrong alarms because of normal fluctuations of the variables. Most frequently, limit checking with fixed thresholds is applied which works well if the process stays in a steady state or if the monitored variable does not depend on the operating point. However, the situation becomes more involved if the monitored variable changes dynamically with other operating points, e.g. forces in rolling mills or machine tools or pressures and temperatures in chemical batch processes.

The advantage of the classical limit-value-based supervision method is their simplicity and reliability for steady-state situations. However, it is only possible to react after a relatively large change of a process feature, i.e. after a large sudden fault or

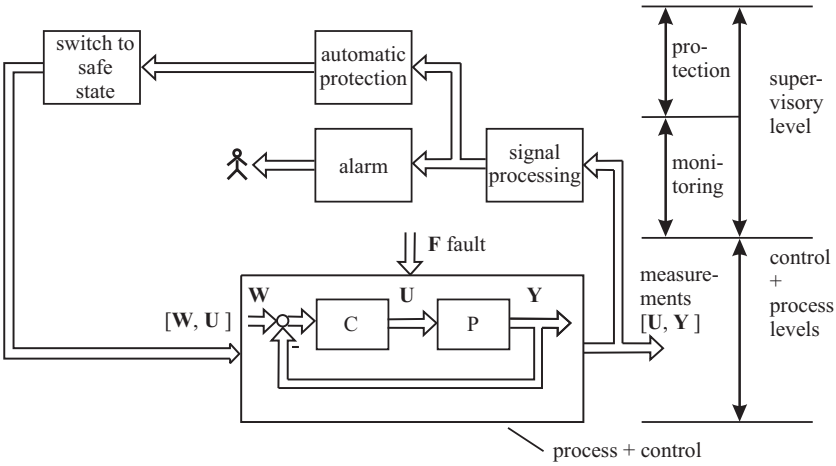


Fig. 2.3. Monitoring and automatic protection

a long-lasting gradually increasing fault. In addition, an in-depth fault diagnosis is usually not possible based on the threshold violation of one or a few variables.

To improve the supervision of technical processes or to improve the quality control of manufactured products a first step could be to implement additional sensors which are related to expected faults and to implement the operator's know-how in computers. However, the use of additional sensors, cables, transmitters, and plugs for getting better information on special faults does not only increase the costs but at the same time deteriorates the overall reliability because the probability of faults increases with more elements. Also the direct software implementation of operator knowledge is a demanding task and does not lead much further without physically based process models.

For large-scale processes with many monitored and limit-checked values, there is another problem: after a severe process fault or failure several alarms may be triggered in a short time, known as "alarm-shower". Consequently, the operators are overloaded with regard to their immediate reactions and to finding the causes of the faulty behavior.

Therefore *advanced methods of supervision, fault detection and fault diagnosis* are required which satisfy the following requirements:

(i) early detection of small faults with abrupt or incipient time behavior
(ii) diagnosis of faults in the processes or process parts and their manipulating devices (actuators) and measurement equipment (sensors)
(iii) detection of faults in closed loops
(iv) supervision of processes in transient states.

The goal for the early fault detection and diagnosis is to have enough time for counteractions such as other operations, reconfiguration, planned maintenance or repair.

Figure 2.4 shows a general scheme that illustrates how, in addition to the classical monitoring and automatic protection, these goals can be reached by automatic means. The intention is to generate more information about the process by using all available measurements and to relate them in the form of mathematical process models. If not only output signals $\mathbf{Y}(t)$ are measured but also the corresponding input signals $\mathbf{U}(t)$, some accessible state variables $\mathbf{x}(t)$ and maybe disturbance signals, then changes of the static and dynamic behavior of the processes by the faults can be used as important information sources. Moreover, also changes of output signals $\Delta\mathbf{Y}(t)$ which are not caused by faults but by input signals $\Delta\mathbf{U}(t)$ or measurable disturbances are automatically taken into account and therefore make the observed comparison variables more sensible to faults. This means that the effects on the outputs $\mathbf{Y}(t)$ by either normal disturbances or faults are automatically separated.

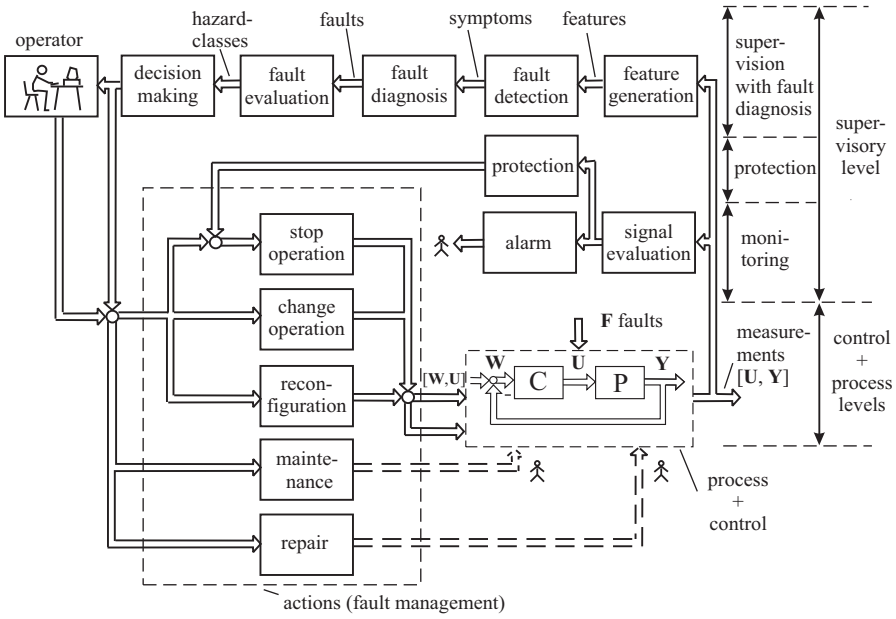


Fig. 2.4. General scheme of advanced supervision methods with fault management (supervisory loop)

The general scheme in Figure 2.4 shows at the third level the following tasks:

3. *Supervision with fault diagnosis*

(a) *feature generation* by, e.g. special signal processing, state estimation, identification and parameter estimation, parity relations or performance measures
(b) *fault detection and generation of symptoms*

(c) *fault diagnosis* by using analytical and also heuristic symptoms and their relations to faults, e.g. by classification methods or reasoning methods via fault-symptom trees. The goal is to determine the kind, size and location of the fault
(d) *fault evaluation* with regard to classify the faults into different hazard classes
(e) *decision on actions* dependent on the hazard class and possible degree of danger. This may be done either automatically or by the operator. Some examples for hazard classes are given in [2.37], Chapter 2.

Based on the gained in-depth information about the condition of the process, further tasks are necessary in order to improve the reliability or safety:

4. *Supervision actions and fault management*: Depending on the hazard classes of the diagnosed fault(s) the following actions can be taken:

(a) *safe operation*, e.g. shut down if there is an imminent danger for the process or the environment
(b) *reliable operation*, e.g. by hindering a further fault expansion through changes of operation state, e.g. operation with lower load, speed, pressure, temperature
(c) *reconfiguration*, e.g. by using other sensors, actuators or redundant (standby) components to keep the process in operation and under control with a "reconfigured" structure
(d) *inspection* to perform a detailed diagnosis by additional measures at the component
(e) *maintenance*, e.g. instantaneously or by the next opportunity to tune process parameters or exchange worn parts
(f) *repair*, e.g. instantaneously to remove a fault or failure or at the next opportunity (overhaul or revision).

These actions are also called *fault management* or *process-oriented asset management* and may incorporate several intermediate actions in the case of redundant systems if the process is in a dangerous state, e.g. for aircraft, power plants, chemical plants or automatic guided vehicles.

Hence, the advanced methods of supervision and following actions are means to improve both the reliability and the safety of technical systems. Of course, these improvements by better information processing and computational intelligence have to be accompanied on the process side by further improving the reliability of all hardware components, by, e.g. proper materials, stress and overall design. Some further interesting developments are:

- *maintenance on demand* (dependent on process condition)
- *tele-diagnosis* with modern communication
- 100% *quality control* of products.

As especially maintenance costs resemble in most cases a high percentage (e.g. $\leq 20\%$) of overall operating costs, the advanced supervision and diagnosis may help to reduce maintenance effort and costs and improve the life time of the processes.

The general scheme in Figure 2.4 shows that there exists a feedback system from faults, signals, features, symptoms, decisions over various actions to compensate for faults. Therefore, this can be called *supervisory loop* or *fault management loop*. However, different to feedback control the signals or states are not all in continuous action. Some parts of information processing like signal evaluation, feature generation and symptom generation may operate continuously, but fault diagnosis, decision making and actions act as discrete events in the case of fault appearance. Hence, the supervisory loop is a *hybrid continuous and discrete event system*.

The known literature on the state of the art of supervision and fault management is mostly related to special processes and distributed in many journals, conference proceedings and books. Some examples are:

- machines: [2.3], [2.19], [2.20], [2.45], [2.60]
- electrical motors: [2.9], [2.24], [2.17], [2.31], [2.67]
- pumps: [2.7], [2.17], [2.21], [2.31], [2.49], [2.64]
- steam turbines: [2.57]
- manufacturing: [2.13], [2.56], [2.65]
- bearings and machinery: [2.8], [2.45], [2.61], [2.68]
- aircraft: [2.47], [2.48], [2.51]
- automotive systems: [2.33], [2.34], [2.44], [2.54]
- chemical processes: [2.23], [2.55].

2.2 Terminology

2.2.1 Faults, failures, malfunctions

As the treated field from faults and failures through reliability, safety and fault-tolerant systems is distributed over many different technological areas, the terminology used is not unique. Various efforts have been made to come to a standardization, for example, the RAM (reliability, availability and maintainability) dictionary, [2.50], in contributions [2.27] and several German standards as DIN and VDI/VDE-Richtlinien (guidelines). The IFAC-Technical Committee SAFEPROCESS has made an effort to come to accepted definitions, [2.38], see also Appendix 13. A survey of related standardization literature is given in the bibliography of Appendix 13. The following sections describe the terminology used in [2.37] and in this book, taking into account the mentioned literature.

Fault:

"A fault is an unpermitted deviation of at least one characteristic property (feature) of the system from the acceptable, usual standard condition."

Remarks:

- a fault is a state within the system

- the unpermitted deviation is the difference between the fault value and the violated threshold of a tolerance zone for its usual value
- a fault is an abnormal condition that may cause a reduction in, or loss of, the capability of a functional unit to perform a required function [2.26]
- there exist many different types of faults, e.g. design fault, manufacturing fault, assembling fault, normal operation fault (e.g. wear), wrong operation fault (e.g. overload), maintenance fault, hardware fault, software fault, operator's fault. (Some of these faults are also called errors, especially if directly caused by humans)
- a fault in the system is independent of whether the system is in operation or not
- a fault may not effect the correct functioning of a system (like a small rent in an axle)
- a fault may initiate a failure or a malfunction
- frequently, faults are difficult to detect, especially if they are small or hidden
- faults may develop abruptly (stepwise) or incipiently (driftwise).

Failure:

"A failure is a permanent interruption of a system's ability to perform a required function under specified operating conditions."

Remarks:

- a failure is the termination of the ability of a functional unit to perform a required function, [2.26]
- a failure is an event
- a failure results from one or more faults
- different types of failures can be distinguished:
  - number of failures: single, multiple
  - predictability:
    - random failure (unpredictable, e.g. statistically independent from operation time or other failures)
    - deterministic failure (predictable for certain conditions)
    - systematic failure or causal failure (dependent on known conditions)
- usually a failure arises after the start of the operation or by increasingly stressing the system.

Malfunction:

"A malfunction is an intermittent irregularity in the fulfillment of a system's desired function."

Remarks:

- a malfunction is a temporary interruption of a system's function

- a malfunction is an event
- a malfunction results from one or more faults
- usually a malfunction arises after the start of the operation or by increasingly stressing the system.

Figure 2.5 shows the relation of faults, failures and malfunctions. The fault may develop abruptly, like a step function, or incipiently, like a drift-like function. The corresponding feature of the system related to the fault is assumed to be proportional to the fault development. After exceeding the tolerance of normal values, the feature indicates a fault at time t_1. Dependent on its size, a failure or a malfunction of the system follows at time t_2.

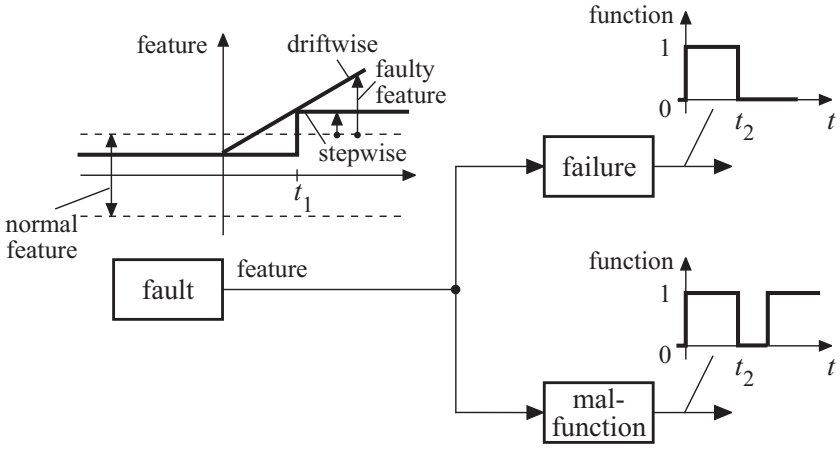


Fig. 2.5. Development of the events "failure" or "malfunction" from a fault which causes a stepwise or driftwise change of a feature

2.2.2 Reliability, availability, safety

With regard to the overall functioning of elements, components, processes and systems the terms reliability, availability and safety play an important role. These terms are considered in more detail in Chapters 3 and 4 of [2.37].

Reliability:

"Ability of a system to perform a required function under stated conditions, within a given scope, during a given period of time."

Remarks:

- short version: ability to perform a required function for a certain period of time

- reliability is quality that lasts over time
- the reliability can be affected by malfunctions and failures
- a measure for reliability is the mean time to failure MTTF $= 1/\lambda$, where λ is the rate of failures per time unit.

Safety:

"Ability of a system not to cause danger to persons or equipment or the environment."

Remarks:

- short version: ability not to cause danger
- safety is concerned with the dangerous effects of faults, failures and malfunctions
- safety can usually be seen as a status, where the risk is not larger than a specified risk limit (risk threshold).

The measures to improve the reliability are oriented towards avoiding faults, failures and malfunctions. Measures for improving safety aim to avoid dangerous effects of failures and malfunctions. An improvement of the reliability generally improves also safety. However, an improvement of safety can result in a deterioration of the reliability if, e.g. the number of components increases. Note that safety and security have similar meanings. *Safety* usually deals with life, equipment or environment, whereas *security* deals with privacy, property, community or state.

Availability:

"Probability that a system or equipment will operate satisfactorily and effectively at any period of time."

Remarks:

- availability is of major importance for the user of a system
- availability takes into account that failures and malfunctions happen and need some time for repair
- a measure for availability is $A = \frac{MTTF}{MTTF+MTTR}$ where MTTR is the mean time to repair
- to reach a high availability MTTF must be large in comparison to MTTR. This can be reached by:
  - large operation time MTTF
    - $\rightarrow$ perfection: highly reliable components
    - $\rightarrow$ tolerance: tolerable faults through redundant structure
  - small repair time MTTR
    - $\rightarrow$ fast and reliable fault diagnosis
    - $\rightarrow$ fast and reliable removal of faults (maintenance repair)
- fault detection and fault diagnosis can improve the availability by early fault detection in combination with maintenance on demand (larger MTTF) and by fast and reliable diagnosis (smaller MTTR).

Dependability:

The term dependability seems not to be clearly defined. Therefore different meanings are cited:

(i) "A form of availability that has the property of always being available when required (and not at any time). It is the degree to which a system is operable and capable of performing its required function at any randomly chosen time during its specific operating time, provided that the system is available at the start of the period." This definition excludes non-operation related influences, [2.50]
(ii) "Dependability is a property of a system that justifies placing one's reliance on it. It covers reliability, availability, safety, maintainability and other issues of importance in critical systems," [2.58].

The [2.26] standard on safety-related systems does not define dependability, only safety integrity.

Integrity:

According to [2.58], the term integrity was earlier defined as:

"The integrity of a system is the ability to detect faults in its own operation and to inform a human operator."

Over the years the meaning was broadened and associated with critical systems. Integrity is frequently used as a synonym for dependability. According to [2.26] it is defined as:

"Safety integrity is the probability of a safety-related system satisfactorily performing the required safety functions under all the stated conditions within a period of time."

Some other expressions like accident, hazard, and risk are defined in [2.37], Chapter 4.

2.2.3 Fault tolerance and redundancy

After applying reliability and safety analysis for the improvement of the design, testing of the product and also corresponding quality control methods during manufacturing, the appearance of certain faults and failures cannot be avoided totally. Therefore, these unavoidable faults should be tolerated by additional design efforts. Hence, high-integrity systems must have the capability *fault tolerance*. This means that faults are compensated in such a way that they do not lead to system failures. After the application of principles to improve the perfection of the components the remaining obvious way to reach this goal is to implement *redundancy*. This means

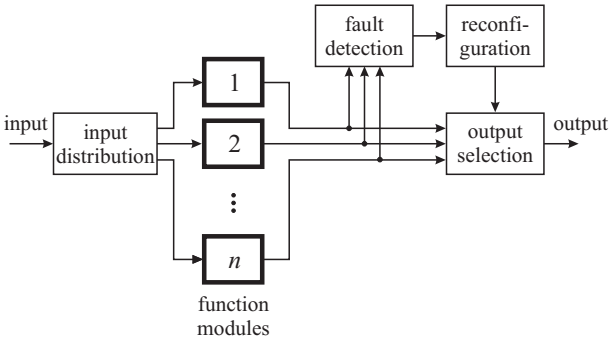


Fig. 2.6. Basic scheme of a fault-tolerant system with parallel function modules as redundance

that in addition to the considered module one or more modules exist as back-up modules, usually in a parallel configuration, see Figure 2.6.

The function modules can be hardware components or software, either identical or diverse. Different arrangements of fault-tolerant systems exist with static or dynamic redundancy, cold or hot standby. In general, the function modules are supervised with fault-detection capability followed by a reconfiguration mechanism to switch off failed modules and to switch on spare modules (dynamic redundancy). The modules are, e.g. actuators, sensors, computers, motors or pumps. For electronic hardware simpler schemes exist with $n \geq 3$ modules and majority voters to build up, e.g. 2-out-of-3 systems (static redundancy). These redundant systems are treated in Part IV.

2.3 Knowledge-based fault detection and diagnosis

As fault detection and fault diagnosis are fundamental for advanced methods of supervision and fault management, these tasks will be considered briefly. Fault detection and diagnosis, in general, are based on measured variables by instruments and observed variables and states by human operators. The automatic processing of measured variables for fault detection requires analytical process knowledge and the evaluation of observed variables requires human expert knowledge which is called *heuristic knowledge*. Therefore fault detection and diagnosis can be considered within a knowledge-based approach, [2.53], [2.59]. Figure 2.7 shows an overall scheme, [2.30], [2.32].

2.3.1 Analytic symptom generation

The analytical knowledge about the process is used to produce quantifiable, analytical information. To do this, data processing based on measured process variables has to be performed to generate first the characteristic values by

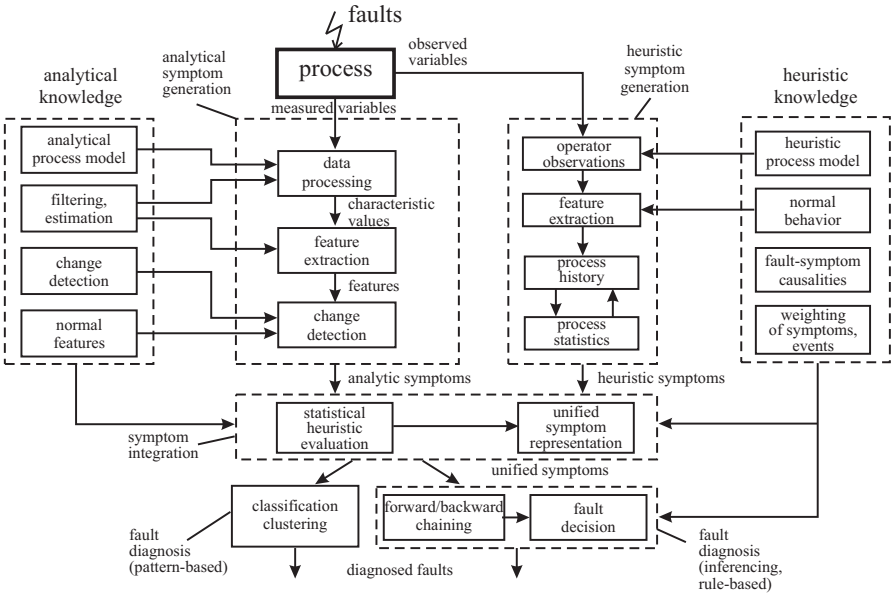


Fig. 2.7. Overall scheme of knowledge-based fault detection and diagnosis

- limit value checking of direct, measurable signals. The characteristic values are the violated signal tolerances
- signal analysis of directly measurable signals by the use of signal models like correlation functions, frequency spectra, autoregressive moving average (ARMA) or the characteristic values, e.g. variances, amplitudes, frequencies or model parameters
- process analysis by using mathematical process models together with parameter estimation, state estimation and parity equation methods. The characteristic values are parameters, state variables or residuals.

In some cases, special features can then be extracted from these characteristic values, e.g. physically defined process coefficients, or special filtered or transformed residuals. These features are then compared with the normal features of the non-faulty process. For this, methods of change detection and classification are applied. The resulting changes (discrepancies) in the mentioned directly measured signals, signal models or process models are considered as analytic symptoms.

Figure 2.8 gives a survey of analytical fault-detection methods. These methods are treated in detail in [2.37] and summarized in the following sections.

2.3.2 Heuristic symptom generation

In addition to the symptom generation using quantifiable information, heuristic symptoms can be produced by using qualitative information from human operators. Through human observation and inspection, heuristic characteristic values in the

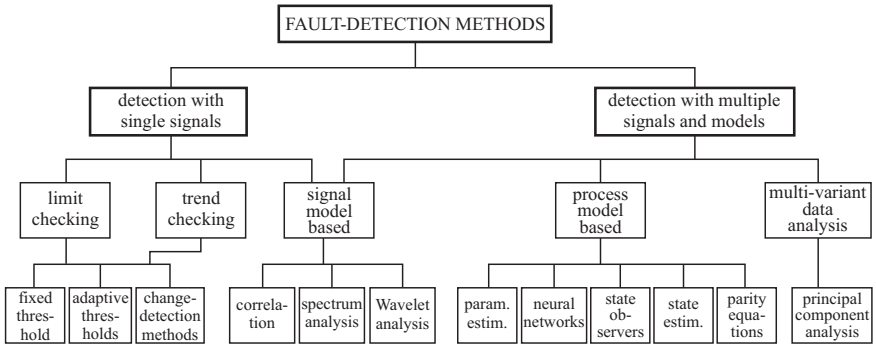


Fig. 2.8. Survey of analytical fault-detection methods

form of special noises, colors, smells, vibration, wear and tear, etc., are obtained. The process history expressed through performed maintenance, repairs, former faults, life-time and load measures, constitutes a further source of heuristic information. Statistical data (e.g. MTTF, fault probabilities) achieved from experience with the same or similar processes can be added. In this way heuristic symptoms are generated, which can be represented as linguistic variables (e.g. small, medium, large) or as vague numbers (e.g. around a certain value).

2.3.3 Fault diagnosis

The task of fault diagnosis consists in determining the type, size and location of the most possible fault, as well as its time of detection.

Fault-diagnosis procedures use the analytic and heuristic symptoms. Therefore they should be presented in an unified form like confidence numbers, membership functions of fuzzy sets or probability density functions after a statistical evaluation over some time. Then either *classification methods* can be applied, if a learned pattern-based procedure is preferred, to determine the faults from symptom patterns or clusters. If, however, more information of fault-symptom relations, e.g. in the form of logic fault-symptom trees or if-then rules are known, *inference methods* (reasoning methods) with forward and backward chaining can be applied.

Figure 2.9 gives a survey of these methods. See also Section 2.6.

2.4 Signal-based fault-detection methods

Fault detection based on single signal measurements is in simple cases performed with limit checking or trend checking, or, in more complex cases, by operating with special signal models, extraction of special signal features and change-detection methods, see Figures 2.8 and 2.10. Corresponding methods are treated in [2.37], Chapters 7 and 8. In the following only a brief description is given.

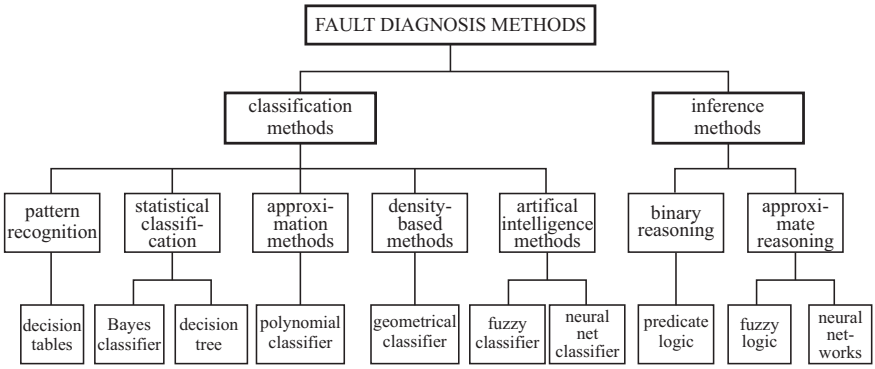


Fig. 2.9. Survey of fault-diagnosis methods

2.4.1 Limit checking of absolute values

Generally, two limit values, called thresholds, are preset, a maximal value Y_{max} and a minimal value Y_{min}. A normal state is when

$$Y_{min} < Y(t) < Y_{max} \tag{2.4.1}$$

which means that the process is in normal situation if the monitored variable stays within a certain tolerance zone. The exceedance of one of the thresholds then indicates a fault somewhere in the process, compare Figure 2.10. This simple method is applied in almost all process automation systems. Examples are the oil pressure (lower limit) or the coolant water (higher limit) of combustion engines, the pressure of the circulation fluid in refrigerators (lower limit) or the control error of a control loop. The thresholds are mostly selected based on experience and represent a compromise. On one hand false alarms through normal fluctuations of the variable should be avoided; on the other, faulty deviations should be detected early. Therefore a trade-off between too narrow and too wide thresholds exists.

2.4.2 Trend checking

A further simple possibility is to calculate the first derivative $\dot{Y} = dY(t)/dt$, the trend of the monitored variable and to check if

$$\dot{Y}_{min} < \dot{Y}(t) < \dot{Y}_{max} \tag{2.4.2}$$

If relatively small thresholds are selected, an alarm can be obtained earlier than for limit checking of the absolute value, see Figure 2.10b). Trend checking is, for example, applied for oil pressures and vibrations of oil bearings of turbines or for wear measures of machines.

Limit checking of absolute values and trends can also be combined, see [2.37].

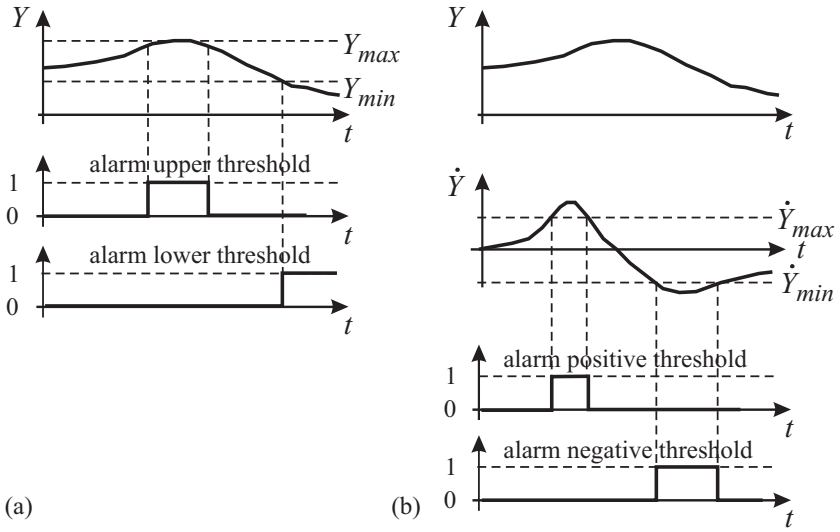


Fig. 2.10. Limit checking: a) absolute value $Y(t)$; b) trend $\dot{Y}(t) = dY(t)/dt$

2.4.3 Change detection with binary thresholds

The monitored variables are usually stochastic variables $Y_i(t)$ with a certain probability density function $p(Y_i)$, mean value and variance

$$\mu_i = E\{Y_i(t)\}; \quad \bar{\sigma}_i^2 = E\left\{[Y_i(t) - \mu_i]^2\right\} \tag{2.4.3}$$

as nominal values for the non-faulty process. Changes are then expressed by

$$\Delta Y_i = E\{Y_i(t) - \mu_i\} \text{ and } \Delta\sigma^2 = E\left\{[\sigma_i(t) - \bar{\sigma}_i]^2\right\} \tag{2.4.4}$$

for $t > t_F$, where t_F is the time of fault occurrence, which is unknown.

If the mean and standard deviations before the change caused by a fault are described by μ_0 and σ_0 and after the change has appeared by μ_1 and σ_1, the change-detection problem is depicted by Figure 2.11, assuming a normal probability distribution of the variable $Y(t)$. Then the following cases of changes can be distinguished:

(i) the mean changes $\mu_1 = \mu_0 + \Delta\mu$; standard deviation $\sigma_1 = \sigma_0$ remains constant
(ii) the mean does not change $\mu_1 = \mu_0$; standard deviation changes $\sigma_1 = \sigma_0 + \Delta\sigma$
(iii) both, mean and standard deviation change.

As an example, case (i) is considered. If the probability densities do not significantly overlap, one can use a fixed threshold.

$$\Delta\, Y_{tol} = \kappa\, \sigma_0 \tag{2.4.5}$$

with, e.g. $\kappa \geq 2$, to detect the change just by observing the average $\mu(Y, t)$. In selecting the threshold, a comparison has to be made between the detection of relatively small changes and false alarms.

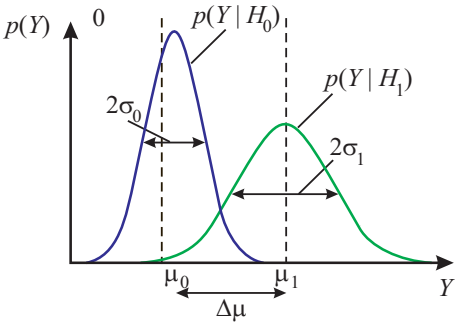


Fig. 2.11. Normal probability density functions of the observed variable Y for the nominal state (index 0) and changed (faulty) state (index 1)

However, the detection problem becomes more involved if the change of the mean

$$\Delta\mu = \mu_1 - \mu_0 \tag{2.4.6}$$

is small compared to the standard deviation, say $\kappa \leq 1$. Then *statistical tests* have to be applied.

The detection of changes of the random variable $Y(k)$ can be performed off-line or on-line in real time. For *off-line change detection* within a sample length N it has to be determined when at some unknown time t_F a change in $Y(k)$ occurred from Y_0 to Y_1. This is only possible after storing all data. For fault detection in *real time* the *on-line change detection* is of more interest. Here at every time k it has to be decided if a change from Y_0 to Y_1 has happened. This means that especially sequential or recursive tests are of interest for fault detection. The first case is easier to decide, because more measurements are available.

Corresponding change-detection methods, taking into account the statistics of the observation, estimation methods, and statistical tests are described in [2.37].

2.4.4 Adaptive thresholds

Process-model-based fault-detection methods described in the next section use process models which do not fully agree with real processes due to model uncertainties. Thus, the generated residuals deviate from zero even without faults. These deviations frequently depend on the amplitude and frequencies of the input excitation. Therefore the residuals may contain a static part which is proportional to the input $U(t)$ and a dynamic part dependent, e.g. on $\dot{U}(t)$. To cope with this problem, [2.25] has introduced an adaptive threshold which uses a first-order high-pass filter (HPF) for enlarging the threshold, Figure 2.12. A proportional enlargement may be added by a constant c_2, [2.15]. A low-pass filter (LPF) is used to smooth the thresholds. The time constants T_1 and T_3 are selected according to the dominating time constant of the process. T_2/T_1 depends on the model uncertainty of the dynamics. Adaptive thresholds were also proposed by [2.5], [2.12].

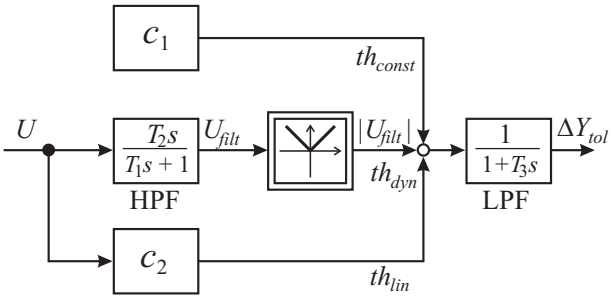


Fig. 2.12. Generation of an adaptive threshold dependent on process input excitation. The constant threshold is $th_{const} = c_1$

2.4.5 Plausibility checks

A rough supervision of measured variables is sometimes performed by checking the plausibility of its indicated values. This means that the measurements are evaluated with regard to credible, convincing values and their compatibility among each other. Therefore, a *single measurement* is examined to find whether the sign is correct and the value is within certain limits. This is also a limit check, however, with usually wide tolerances. If *several measurements* are available for the same process then the measurements can be related to each other with regard to their normal ranges by using logic rules, like

$$\text{IF } [Y_{1min} < Y_1(t) < Y_{1max}] \text{ THEN } [Y_{2min} < Y_2(t) < Y_{2max}] \tag{2.4.7}$$

For example, one expects for a circulation pump with rotating speed n and pressure p

$$\text{IF } [1000 \text{ rpm} < n < 3000 \text{ rpm}] \text{ THEN } [3 \text{ bar} < p < 8 \text{ bar}]$$

The plausibility check can also be made dependent on the operating condition, like

$$\text{IF } [\text{Operating condition 1}] \text{ THEN } [Y_{3min} < Y_3(t) < Y_{3max}] \tag{2.4.8}$$

One example is the oil pressure p_{oil} of a combustion engine with speed n and cooling water temperature ϑ_{H20}:

$$\text{IF } [n < 1500\,\text{rpm}] \text{ AND } [\vartheta_{\text{H20}} < 50^{\circ}\text{C}] \text{ THEN } [3\,\text{bar} < p_{oil} < 5\,\text{bar}] \tag{2.4.9}$$

Hence, plausibility checks may be formulated by using rules with binary logic connections like AND, OR. These rules and ranges of the measurements allow a rough description of the expected behavior of the process under normal conditions. If these rules are not satisfied either the process or the measurements are faulty. Then, one needs further testing to localize the fault and its cause.

These plausibility checks presuppose the ranges of measured process variables under certain operating conditions and represent rough process models. If the ranges

of the variables are increasingly made smaller, many rules would be required to describe the process behavior. Then, it is better to use mathematical process models in the form of equations to detect abnormalities. Therefore, plausibility tests can be seen as a first step towards model-based fault-detection methods.

2.4.6 Signal-analysis methods

Many measured signals of processes show oscillations that are either of harmonic or stochastic nature, or both. If changes of these signals are related to faults in the actuators, the process and sensors, signal-model-based fault-detection methods can be applied. Especially for machine vibration, the measurement of position, speed or acceleration allows one to detect, for example, imbalance or bearing faults (turbo machines), knocking (gasoline engines) and chattering (metal grinding machines). But also signals from many other sensors, like electrical current, position, speed, force, flow and pressure, frequently contain oscillations with a variety of higher frequencies than the process dynamics.

The task of fault detection by the analysis of signal models is summarized in Figure 2.13. By assuming special mathematical models for the measured signal, suitable features are calculated, for example, amplitudes, phases, spectrum frequencies and correlation functions for a certain frequency band width $\omega_{min} \leq \omega \leq \omega_{max}$ of the signal. A comparison with the observed features for normal behavior provides changes of the features which then are considered as *analytical symptoms*.

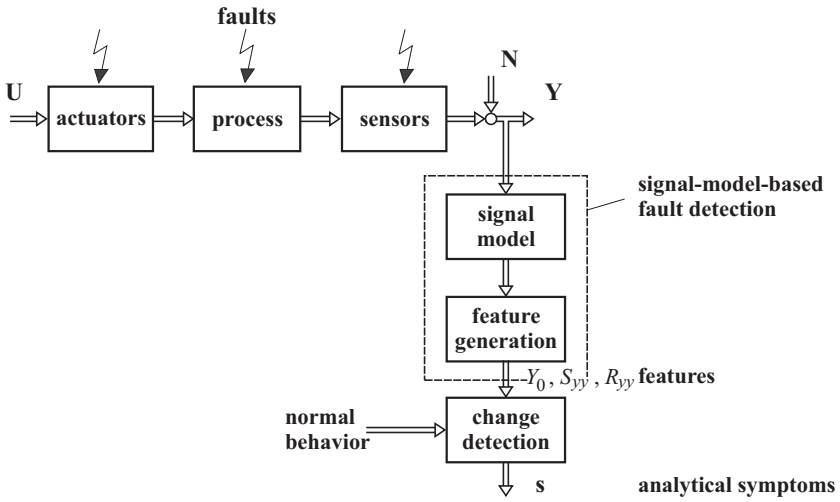


Fig. 2.13. Scheme for the fault detection with signal models

The signal models can be divided into nonparametric models, like frequency spectra or correlation functions, or parametric models, like amplitudes for distinct

frequencies or ARMA-type models. Signal-analysis methods exist for harmonic oscillations, stochastic signals and non-stationary signals, compare the scheme of Figure 2.14.

For the analysis of stationary periodic signals band pass filtering or Fourier analysis can be used. Non-stationary periodic signals may be analyzed with, e.g. wavelet transforms. The analysis of stochastic signals is frequently performed by correlation functions, spectrum analysis and signal parameter estimation for ARMA-models. These methods are treated in [2.37], where also many other references are given.

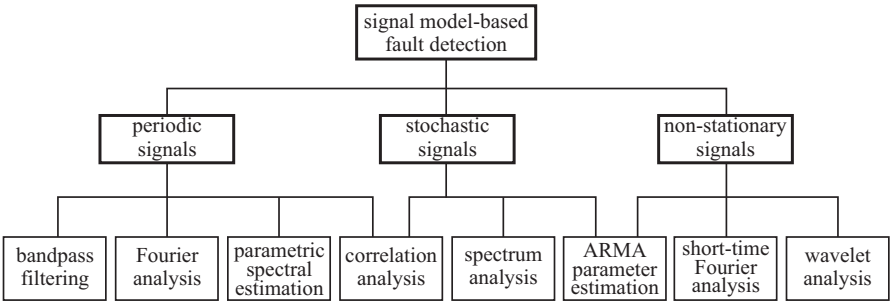


Fig. 2.14. Survey of signal-analysis methods for signal-model-based fault detection

2.5 Process-model-based fault-detection methods[2]

Different approaches for fault detection using mathematical models have been developed in the last few decades (see, e.g. [2.66], [2.22], [2.28], [2.39], [2.32], [2.18], [2.11], [2.4], [2.52]). The task consists of the detection of faults in the processes, actuators and sensors by using the *dependencies between different measurable signals*. These dependencies are expressed by mathematical process models. Figure 2.15 shows the basic structure of model-based fault detection. Based on measured input signals $\mathbf{U}$ and output signals $\mathbf{Y}$, the detection methods generate residuals $\mathbf{r}$, parameter estimates $\hat{\boldsymbol{\Theta}}$ or state estimates $\hat{\mathbf{x}}$, which are called features. By comparison with the normal features (nominal values), changes of features are detected, leading to analytical symptoms $\mathbf{s}$.

For the application of model-based fault-detection methods, the process configurations according to Figure 2.16 have to be distinguished. With regard to the inherent dependencies used for fault detection, and the possibilities for distinguishing between different faults, the situation improves greatly from case a) to b) or c) or d), by the availability of some more measurements.

[2] This chapter follows [2.36]

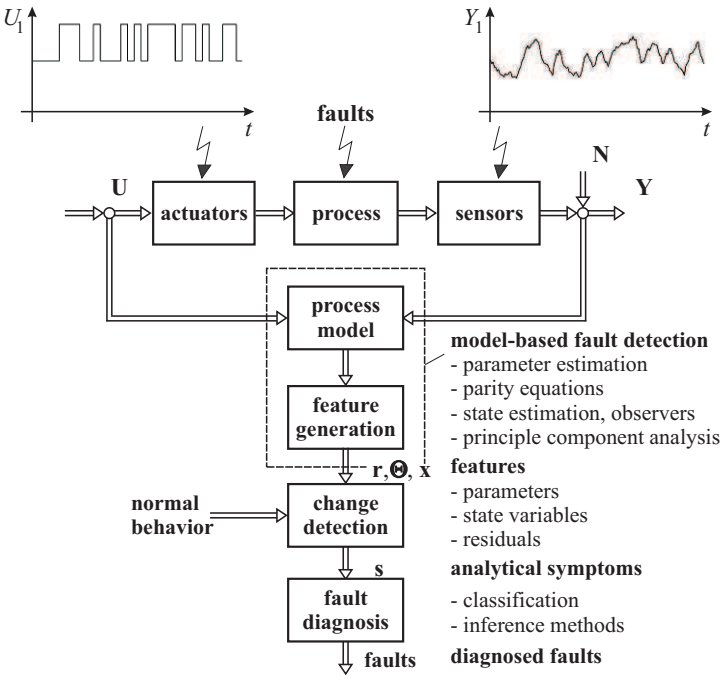


Fig. 2.15. General scheme of process-model-based fault detection and diagnosis

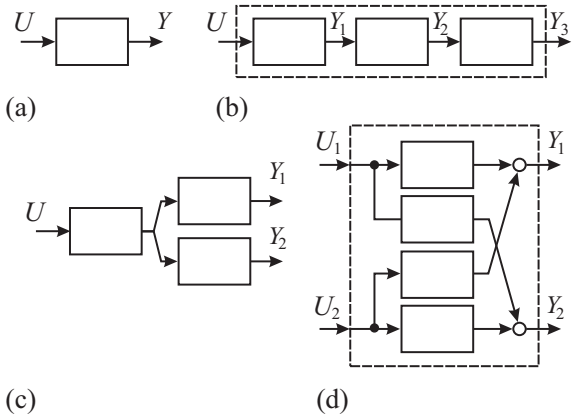


Fig. 2.16. Process configuration for model-based fault detection: a) SISO (single-input single-output); b) SISO with intermediate measurements; c) SIMO (single-input multi-output); d) MIMO (multi-input multi-output)

2.5.1 Process models and fault modeling

A fault is defined as an unpermitted deviation of at least one characteristic property of a variable from an acceptable behavior. Therefore, the fault is a state that may lead to a malfunction or failure of the system. The time dependency of faults can be distinguished, see Figure 2.17, as abrupt fault (stepwise), incipient fault (drift-like), or intermittent fault. With regard to the process models, the faults can be further classified. According to Figure 2.18 *additive faults* influence a variable Y by an addition of the fault f, and *multiplicative faults* by the product of another variable U with f. Additive faults appear, e.g., as offsets of sensors, whereas multiplicative faults are parameter changes within a process.

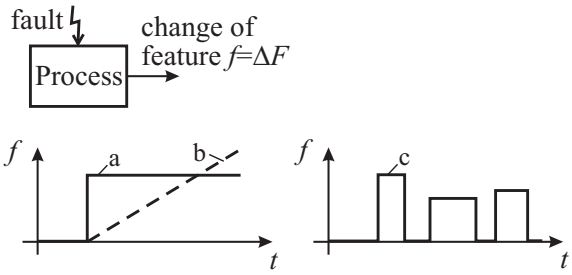


Fig. 2.17. Time-dependency of faults: a) abrupt; b) incipient; c) intermittent

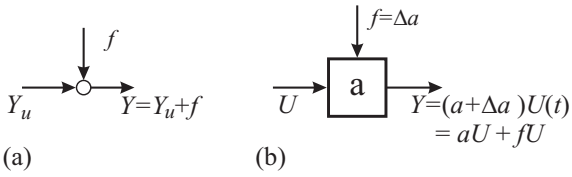
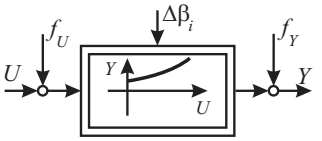


Fig. 2.18. Basic models of faults: a) additive fault; b) multiplicative faults

Now lumped-parameter processes are considered, which operate in open loop. The *static behavior* (steady states) is frequently expressed by a nonlinear characteristic as shown in Table 2.1. Changes of parameters β_i can be obtained by parameter estimation with, e.g., methods of least squares, based on measurements of different input output pairs $[Y_j, U_j]$. This method is applicable for, e.g. valves, pumps, drives, and engines.

More information on the process can usually be obtained with dynamic process models. Table 2.2 shows the basic input/output models in the form of a differential equation or a state-space model as vector differential equation. Similar representations hold for nonlinear processes and for multi-input multi-output processes, also in discrete time.

Table 2.1. Fault detection of a nonlinear static process via parameter estimation for steady states



Measured signals: $U(t)$, $Y(t)$

Basic equation:

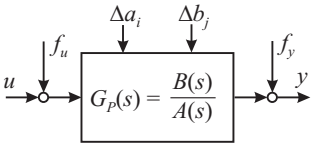
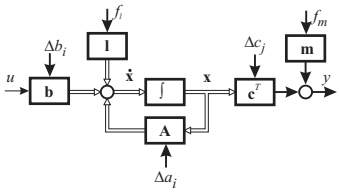
$$Y = \beta_0 + \beta_1 U + \beta_2 U^2 + ... + \beta_q U^q \rightarrow Y = \boldsymbol{\psi}_S^T \boldsymbol{\Theta}_S$$

$$\boldsymbol{\Theta}_S^T = [\beta_0 \beta_1 \ldots \beta_q] \qquad \boldsymbol{\psi}_S^T = [1 \; U \; U^2 \ldots U^q]$$

Additive faults: f_U input fault; f_Y output fault

Multiplicative faults: $\Delta\beta_i$ parameter faults

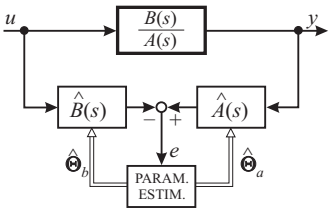
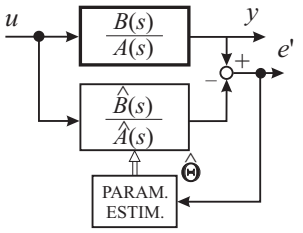
Table 2.2. Linear dynamic process models and fault modeling

Input/output model	State-space model
 Measured signals: $y(t) = Y(t) - Y_{00}$; $u(t) = U(t) - U_{00}$	
Basic equations: $y(t) + a_1 y^{(1)}(t) + ... + a_n y^{(n)}(t) = b_0 u(t) + b_1 u^{(1)}(t) + ... + b_m u^{(m)}(t)$	$\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{b}u(t)$ $y(t) = \mathbf{c}^T \mathbf{x}(t)$
$y(t) = \boldsymbol{\psi}^T(t)\,\boldsymbol{\Theta}$ $\boldsymbol{\Theta}^T = [\,a_1 \ldots a_n \; b_0 \ldots b_m\,]$ $\boldsymbol{\psi}^T = [\,-y^{(1)}(t) \ldots -y^{(n)}(t) \; u(t) \ldots u^{(m)}(t)\,]$	$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_1 \\ 1 & 0 & -a_2 \\ \vdots & \vdots & \vdots \end{bmatrix}$ $\mathbf{b}^T = [\,b_0 \quad b_1 \quad \cdots\,]$ $\mathbf{c}^T = [\,0 \quad 0 \quad \cdots \quad 1\,]$
Additive faults: f_u input fault; f_y output fault	f_l input or state-variable fault f_m output fault
Multiplicative faults: Δa_i, Δb_j parameter faults	$\Delta\mathbf{A}$, $\Delta\mathbf{b}$, $\Delta\mathbf{c}$ parameter faults

2.5.2 Fault detection with parameter estimation

Process-model-based methods require the knowledge of a usually dynamic process model in the form of a mathematical structure and parameters. For linear processes in continuous time the models can be impulse responses (weighting functions), differential equations of frequency responses. Corresponding models for discrete-time (after sampling) are impulse responses, difference equations or z-transfer functions. For fault detection in general, differential equations or difference equations are primarily suitable. In most practical cases the process parameters are partially not known or not known at all. Then, they can be determined with parameter estimation methods by measuring input and output signals if the basic model structure is known. Table 2.3 shows two approaches by minimization of the equation error and the output error. The first one is linear in the parameters and allows therefore direct estimation of the parameters (least-squares estimates) in non-recursive or recursive form. The second one needs numerical optimization methods and therefore iterative procedures, but may be more precise under the influence of process disturbances. The symptoms are deviations of the process parameters $\Delta\boldsymbol{\Theta}$. As the process parameters $\boldsymbol{\Theta} = f(\mathbf{p})$ depend on physically defined process coefficients $\mathbf{p}$ (like stiffness, damping coefficients, resistance), determination of changes $\Delta\mathbf{p}$ allows usually a deeper insight and makes fault diagnosis easier, [2.29]. Parameter estimation methods operate with adaptive process models, where only the model structure is known. They usually need a dynamic process input excitation and are especially suitable for the detection of multiplicative faults.

Table 2.3. Fault detection with parameter estimation methods for dynamic processes

Minimization of equation error	Minimization of output error
	
Loss function: $V = \Sigma\, e^2(k)$ Method: • non-recursive $\hat{\boldsymbol{\Theta}} = [\boldsymbol{\Psi}^T\boldsymbol{\Psi}]^{-1}\boldsymbol{\Psi}^T\mathbf{y}$ • recursive $\hat{\boldsymbol{\Theta}}(k+1) = \hat{\boldsymbol{\Theta}}(k) + \boldsymbol{\gamma}(k)e(k+1)$	$V = \Sigma\, e'^2(k)$ • nonlinear parameter optimization • recursive form $\hat{\boldsymbol{\Theta}}(\nu+1) = \hat{\boldsymbol{\Theta}}(\nu) + \boldsymbol{\Gamma}(\nu)\frac{\partial V}{\partial \boldsymbol{\Theta}}(\nu)$
Symptoms: • model parameters $\Delta\hat{\boldsymbol{\Theta}}(j) = \hat{\boldsymbol{\Theta}}(j) - \boldsymbol{\Theta}_0$ • process coefficients $\hat{\mathbf{p}} = f^{-1}[\boldsymbol{\Theta}]$ $\Delta\mathbf{p}(j) = \hat{\mathbf{p}}(j) - \mathbf{p}_0$	

2.5.3 Fault detection with state observers and state estimation

If the process parameters are known, either state observers or output observers can be applied, Table 2.4. Fault modeling is then performed with additive faults $\mathbf{f}_L$ at the input (additive actuator or process faults) and $\mathbf{f}_M$ at the output (sensor offset faults).

a) State observers

The classical state observer can be applied if the faults can be modeled as state-variable changes $\Delta\mathbf{x}_i$ e.g., for leaks. Special design of the matrix $\mathbf{W}$ allows one to generate structured residuals. In the case of *multi-output processes* special arrangements of observers were proposed:

Dedicated observers for multi-output processes

- Observer, excited by one output: One observer is driven by one sensor output. The other outputs $\mathbf{y}$ are reconstructed and compared with measured outputs $\mathbf{y}$. This allows the detection of single sensor faults, [2.6]
- Bank of observers, excited by all outputs: Several state observers are designed for a definite fault signal and detected by a hypothesis test, [2.66]
- Bank of observers, excited by single outputs: Several observers for single sensor outputs are used. The estimated outputs $\mathbf{y}$ are compared with the measured outputs $\mathbf{y}$. This allows the detection of multiple sensor faults, [2.6] (dedicated observer scheme)
- Bank of observers, excited by all outputs except one: As before, but each observer is excited by all outputs except one sensor output which is supervised, [2.10].

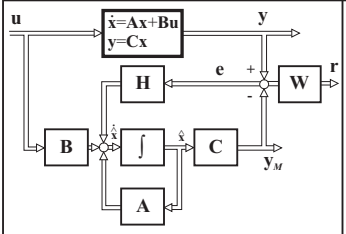
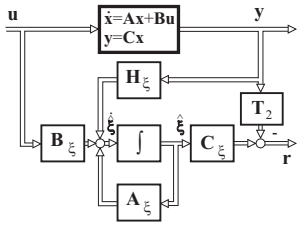
Fault-detection filters (fault-sensitive filters) for multi-output processes

The feedback $\mathbf{H}$ of the state observer is chosen so that particular fault signals $\mathbf{f}_L(t)$ change in a definite direction and fault signals $\mathbf{f}_M(t)$ in a definite plane, [2.2] and [2.43].

b) Output observers

Another possibility is the use of output observers (or unknown input observers) if the reconstruction of the state variables $\mathbf{x}(t)$ is not of interest. A linear transformation then leads to new state variables $\boldsymbol{\xi}(t)$. The residuals $\mathbf{r}(t)$ can be designed such that they are independent of the unknown inputs $\mathbf{v}(t)$, and of the state by special determination of the matrices $\mathbf{C}_\xi$ and $\mathbf{T}_2$. The residuals then depend only on the additive faults $\mathbf{f}_L(t)$ and $\mathbf{f}_M(t)$. However, all process model matrices must be known precisely. Hence, the observer-based fault-detection methods operate with a *fixed parameter model* and correct the state variables by the feedback of output errors. A comparison with the parity equation approach shows similarities.

Table 2.4. Fault detection with observers for dynamic processes

State observer	Output observer
Process model: $\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) + \mathbf{F}\,\mathbf{v}(t) + \mathbf{L}\,\mathbf{f}_L(t)$ $\mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t) + \mathbf{N}\,\mathbf{n}(t) + \mathbf{M}\,\mathbf{f}_M(t)$ $\mathbf{v}(t)$, $\mathbf{n}(t)$: disturbance signals; $\mathbf{f}_L$, $\mathbf{f}_M$: additive fault signals	
	
Observer equations: $\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}\,\hat{\mathbf{x}}(t) + \mathbf{B}\,\mathbf{u}(t) + \mathbf{H}\,\mathbf{e}(t)$ $\mathbf{e}(t) = \mathbf{y}(t) - \mathbf{C}\,\hat{\mathbf{x}}(t)$	$\dot{\hat{\boldsymbol{\xi}}}(t) = \mathbf{A}_\xi\,\hat{\boldsymbol{\xi}}(t) + \mathbf{B}_\xi\,\mathbf{u}(t) + \mathbf{H}_\xi\,\mathbf{y}(t)$ $\boldsymbol{\eta}(t) = \mathbf{C}_\xi\,\hat{\boldsymbol{\xi}}(t)$ $\xi(t) = \mathbf{T}_1\,\mathbf{x}(t)$: transformation
Residuals: • $\Delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}_0(t)$ • $\mathbf{e}(t)$ • $\mathbf{r}(t) = \mathbf{W}\,\mathbf{e}(t)$ Special observers: – fault-sensitive filters (**H** such that $\mathbf{r}(t)$ defin. direct.) – dedicated observers (for different sensor outputs)	$\boldsymbol{\xi}(t) = \hat{\boldsymbol{\xi}}(t) - \mathbf{T}_1\,\hat{\mathbf{x}}(t)$ $\mathbf{r}(t) = \mathbf{C}_\xi\,\boldsymbol{\xi}(t) - \mathbf{T}_2\mathbf{M}f_M(t)$ – independent of $\mathbf{x}(t)$, $\mathbf{u}(t)$, $\mathbf{v}(t)$ – dependent on $\mathbf{f}_L(t)$, $\mathbf{f}_M(t)$ Design equations: $\mathbf{T}_1\mathbf{A} - \mathbf{A}_\xi\mathbf{T}_1 = \mathbf{H}_\xi\mathbf{C}$ $\mathbf{B}_\xi = \mathbf{T}_1\mathbf{B}$ $\mathbf{T}_1\mathbf{V} = \mathbf{0}$ $\mathbf{C}_\xi\mathbf{T}_1 - \mathbf{T}_2\mathbf{C} = \mathbf{0}$

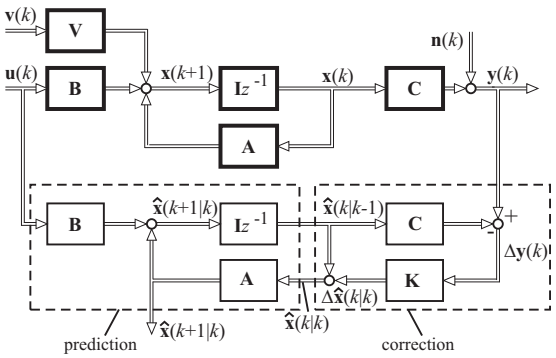
c) State estimation

Whereas state observers are designed for deterministic initial states $\mathbf{x}(0)$ and inputs $\mathbf{u}$ and no disturbances, state estimators are optimized filters for stochastic initial states, stochastic state disturbances $\mathbf{v}$ at the input and stochastic disturbances $\mathbf{n}$ at the output with known covariances. In the case of continuous-time signals the Kalman–Bucy filter results and for discrete-time signals the Kalman filter. Table 2.5 shows the signal flow and basic equations for the mostly used Kalman filter with discrete-time signals, see also [2.37], and the cited references.

The application of the Kalman filter is similar to that of state observers and should only be applied if considerable stochastic disturbances act on the input and/or the output signals. However, the covariance matrices of both disturbances must be known for determining the filter gain $\bar{\mathbf{K}}$. This needs in many cases some trials to find appropriate values.

Table 2.5. Fault detection with state estimation and discrete-time signals (Kalman filter)

State estimation (Kalman filter)



Process models:

$\mathbf{x}(k+1) = \mathbf{A}\,\mathbf{x}(k) + \mathbf{B}\,\mathbf{u}(k) + \mathbf{V}\,\mathbf{v}(k)$

$\mathbf{y}(k) = \mathbf{C}\,\mathbf{x}(k) + \mathbf{n}(k)$

$\mathbf{v}(k)$, $\mathbf{n}(k)$: stochastic disturbances with known covariance matrices $\mathbf{M}$ and $\mathbf{N}$

State estimation equations:

- prediction: $\hat{\mathbf{x}}(k+1|k) = \mathbf{A}\,\hat{\mathbf{x}}(k|k) + \mathbf{B}\,\mathbf{u}(k)$
- correction: $\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k\text{-}1) + \overline{\mathbf{K}}[\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k|k-1)]$
- filter gain: $\overline{\mathbf{K}} = \mathbf{P}^{-}\mathbf{C}^{\mathrm{T}}[\mathbf{C}\,\mathbf{P}^{-}\mathbf{C}^{\mathrm{T}} + \mathbf{N}]^{-1}$
- error covariance matrix Riccati equation:

$\mathbf{P}^{-}(k+1) = \mathbf{A}\mathbf{P}^{-}(k)\,\mathbf{A}^{\mathrm{T}} - \mathbf{A}\mathbf{P}^{-}(k)\,\mathbf{C}^{\mathrm{T}}[\mathbf{C}\mathbf{P}^{-}(k)\,\mathbf{C}^{\mathrm{T}} + \mathbf{N}]^{-1}\mathbf{C}\mathbf{P}^{-}(k)^{\mathrm{T}}\mathbf{A} + \mathbf{V}\mathbf{M}\,\mathbf{V}^{\mathrm{T}}$

Residuals: $\Delta\hat{\mathbf{x}}(k+1|k)$

$\Delta\mathbf{y}(k) = \mathbf{e}\,(k)$

2.5.4 Fault detection with parity equations

A straightforward model-based method of fault detection is to take a fixed model G_M and run it parallel to the process, thereby forming an output error, see Table 2.6:

$$r'(s) = [G_p(s) - G_M(s)]\,u(s) \tag{2.5.1}$$

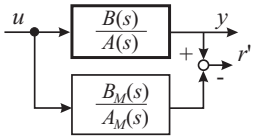
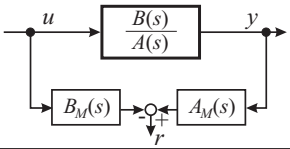
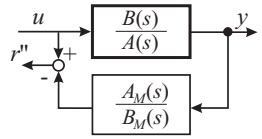
If $G_p(s) = G_M(s)$, the output error for additive input and output faults becomes, Table 2.2:

$$r'(s) = G_p(s)\,f_u(s) + f_y(s) \tag{2.5.2}$$

Another possibility is to generate an equation error (polynominal error) or an input error as in Table 2.7, [2.18].

In all cases, the residuals only depend on the additive input faults $f_u(t)$ and output faults $f_y(t)$. The same procedure can be applied for multivariable processes by using a state-space model, see Table 2.7.

Table 2.6. Fault detection with different forms of parity equations for linear input/output models

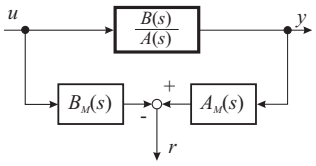
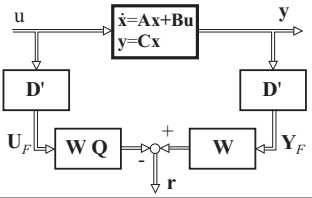
Output error	Equation error	Input error
		
Parity equations: $r'(s) = y(s) - \frac{B_M(s)}{A_M(s)}\, u(s)$ $r'(t) = \boldsymbol{\psi}_r'^T(t)\boldsymbol{\Theta}_{Mr} + \boldsymbol{\psi}_a^T(t)\boldsymbol{\Theta}_{Ma} - \boldsymbol{\psi}_b^T(t)\boldsymbol{\Theta}_{Mb}$	$r(s) = A_M(s)\, y(s) - B_M\, u(s)$ $r(t) = \boldsymbol{\psi}_a^T(t)\, \boldsymbol{\Theta}_{Ma} - \boldsymbol{\psi}_b^T \boldsymbol{\Theta}_{Mb}$	$r''(s) = u(s) - \left(\frac{A_M(s)}{B_M(s)}\right) y(s)$ $r''(t) = \boldsymbol{\psi}_b''^T(t)\boldsymbol{\Theta}_{Mb}'' - \boldsymbol{\psi}_a^T(t)\boldsymbol{\Theta}_{Ma}$
$B_M(s) = b_0 + b_1 s + \ldots + b_m s^m$ $A_M(s) = 1 + a_1 s + \ldots + a_n s^n$ $\boldsymbol{\Theta}_{Mr}^T = [\,1\ a_1\, a_2 \ldots\ a_n\,]$ $\boldsymbol{\psi}_r'^T = [\,y\ y^{(1)}\, y^{(2)} \ldots\ y^{(n)}\,]$	$\boldsymbol{\Theta}_{Mb}^T = [\,b_0\, b_1 \ldots\ b_m\,]$ $\boldsymbol{\Theta}_{Ma}^T = [\,1\ a_1 \ldots\ a_n\,]$ $\boldsymbol{\psi}_b^T = [\,u\ u^{(1)} u^{(2)} \ldots\ u^{(m)}\,]$ $\boldsymbol{\psi}_a^T = [\,y\ y^{(1)} y^{(2)} \ldots\ y^{(n)}\,]$	$\boldsymbol{\Theta}_{Mb}''^T = \frac{1}{b_0}[\,1\ b_1\, b_2 \ldots\ b_m\,]$ $\boldsymbol{\psi}_b''^T = [\,u\ u^{(1)} u^{(2)} \ldots\ u^{(m)}\,]$

The derivatives of the signals can be obtained by state-variable filters, [2.24]. Corresponding equations exist for discrete time and are easier to implement for the state-space model. The residuals shown in Table 2.6 and 2.7 left are direct residuals. If the parity equations are formulated for more than one input and one output, it becomes possible to generate structured residuals such that faults do not influence all residuals. This improves the isolability of faults, [2.18]. For example, the components of matrix $\mathbf{W}$ for the state-space model, Table 2.7 right, are selected such that, e.g., one measured variable has no impact on a specific residual. Parity equations are suitable for the detection of additive faults. They are simpler to design and to implement than output observer-based approaches and lead approximately to the same results. A comparison of fault detection with observers, Kalman filter and parity equations is given in [2.37], Section 11.4.

2.5.5 Direct reconstruction of non-measurable variables

State observers and Kalman filters reconstruct non-measurable variables contained in the state-vector $\mathbf{x}(t)$ and parameter-estimation methods reconstruct non-measurable parameters $\boldsymbol{\Theta}$ from measured input signals $u(t)$ and output signals $y(t)$. However, process models or parts of it can also be directly used to calculate non-measurable variables from measured variables, for example by using algebraic relationships. A first example is the calculation of the torque M of a DC motor from the current I by using $M(t) = \Psi I(t)$, where Ψ is the flux linkage. A second example is the reconstruction of the volume flow rate $\dot{V}$ of a centrifugal pump transporting a liquid through a pipe from the rotor angular speed ω by $\dot{V}(t) = \kappa\omega(t)$ for steady-state operation, [2.67]. This kind of reconstruction via algebraic relations holds especially

Table 2.7. Fault detection with parity equations for dynamic processes

Input/output model, equation error	State space model
	
Parity equations: $r(s) = A_M(s)y(s) - B_M(s)u(s)$ $r(t) = \boldsymbol{\psi}_a^T(t)\boldsymbol{\Theta}_{Ma} - \boldsymbol{\psi}_b^T(t)\boldsymbol{\Theta}_{Mb}$	$\mathbf{Y}_F(t) = \mathbf{T}\,\mathbf{X}(t) + \mathbf{Q}\,\mathbf{U}_F(t)$ $\mathbf{W}\,\mathbf{Y}_F(t) = \mathbf{W}\,\mathbf{T}\,\mathbf{x}(t) + \mathbf{W}\,\mathbf{Q}\,\mathbf{U}_F(t)$ $\mathbf{W}\,\mathbf{T} = 0$ $\mathbf{r}(t) = \mathbf{W}(\mathbf{Y}_F(t) - \mathbf{Q}\,\mathbf{U}_F(t))$
$B_M(s) = b_0 + b_1 s + \ldots + b_m s^m$ $A_M(s) = 1 + a_1 s + \ldots + a_n s^n$ $\boldsymbol{\Theta}_{Mb}^T = [\,b_0\, b_1 \ldots\, b_m\,]$ $\boldsymbol{\Theta}_{Ma}^T = [\,1\, a_1\, a_2 \ldots\, a_n\,]$ $\boldsymbol{\psi}_b^T = [\,u\, u^{(1)} u^{(2)} \ldots\, u^{(m)}\,]$ $\boldsymbol{\psi}_a^T = [\,y\, y^{(1)} \ldots\, y^{(n)}\,]$	$\mathbf{D'u} = [\,u\, u^{(1)} \ldots\, u^{(m)}\,]^T = \mathbf{U}_F$ $\mathbf{D'y} = [\,\mathbf{y}\, \mathbf{y}^{(1)} \ldots\, \mathbf{y}^{(n)}\,]^T = \mathbf{Y}_F$ $\mathbf{T} = [\,C\; CA\; CA^2 \ldots\,]^T$ $\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 \ldots. \\ CB & 0 & 0 \\ CAB & CB & 0 \\ \mathrm{M} & & \end{bmatrix}$

for transformers and converters whose behavior is expressed with the power covariables effort $e(t)$ and flow $f(t)$, [2.35].

2.6 Fault-diagnosis methods

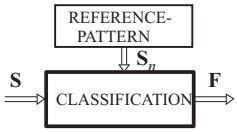
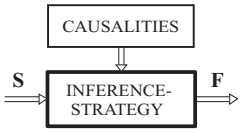
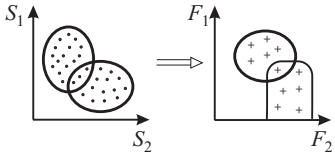
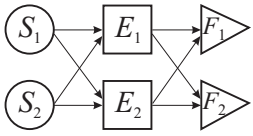
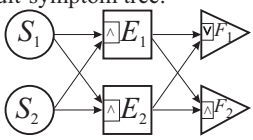
The task of fault diagnosis consists of the determination of the type of fault with as many details as possible such as the fault size, location and time of detection. The diagnostic procedure is based on the observed analytical and heuristic symptoms and the heuristic knowledge of the process, see the schemes in Figures 2.7, 2.9 and 2.15. The inputs to a knowledge-based fault-diagnosis system are all available symptoms as facts and the fault-relevant knowledge about the process, mostly in heuristic form. The symptoms may be presented just as binary values [0, 1] or, e.g., fuzzy sets to take gradual sizes into account.

2.6.1 Classification methods

If no further knowledge is available for the relations between features and faults classification or pattern recognition methods can be used, Table 2.8. Here, reference vectors $\mathbf{S}_n$ are determined for the normal behavior. Then the corresponding input

vectors **S** of the symptoms are determined experimentally for certain faults F_j applying the fault-detection methods. The relationship between **F** and **S** is therefore learned (or trained) experimentally and stored, forming an *explicit knowledge base*. By comparison of the observed **S** with the normal reference $\mathbf{S}_n$, faults **F** can be concluded.

Table 2.8. Methods of fault diagnosis

Classification methods	Inference methods
	
Without a-priori knowledge on symptom causalities Mapping:  $\mathbf{S}^T = [\, S_1, S_2 \ldots S_n \,]$ $\mathbf{F}^T = [\, F_1, F_2 \ldots F_m \,]$	With a-priori knowledge on symptom causalities Causal network:  Fault-symptom tree: 
Classification: – statistical – geometrical – neural nets – fuzzy clusters	Rules: If $< S_1 \wedge S_2 >$ Then $< E_1 >$ Diagnostic reasoning: – Boolean logic: facts binary – Approximative reasoning: – Probabilistic facts: probability densities – Fuzzy facts: fuzzy sets

One distinguishes between *statistical* or *geometrical classification methods*, with or without certain probability functions, [2.63]. A further possibility is the use of neural networks because of their ability to approximate nonlinear relations and to determine flexible decision regions for **F** in continuous or discrete form, [2.46]. By fuzzy clustering the use of fuzzy separation areas is possible.

2.6.2 Inference methods

For some technical processes, the basic relationships between faults and symptoms are at least partially known. Then this a-priori knowledge can be represented in causal relations: fault → events → symptoms. Table 2.8 shows a simple causal network, with the nodes as states and edges as relations. The establishment of these causalities follows the fault-tree analysis (FTA), proceeding from faults through intermediate events to symptoms (the physical causalities) or the event-tree analysis (ETA), proceeding from the symptoms to the faults (the diagnostic forward-chaining causalities). To perform a diagnosis, this qualitative knowledge can now be expressed in the form of rules: IF < condition > THEN < conclusion >. The condition part (premise) contains facts in the form of symptoms S_i as inputs, and the conclusion part includes events E_k and faults F_j as a logical cause of the facts. If several symptoms indicate an event or fault, the facts are associated by AND and OR connectives, leading to rules in the form

$$\begin{aligned} &\text{IF} \quad < S_1 \quad \text{AND} \quad S_2 > \text{THEN} \quad < E_1 > \\ &\text{IF} \quad < E_1 \quad \text{OR} \quad E_2 > \text{THEN} \quad < F_1 > . \end{aligned}$$

For the establishment of this heuristic knowledge several approaches exist, see [2.14], [2.62]. In the classical fault-tree analysis the symptoms and events are considered as binary variables, and the condition part of the rules can be calculated by Boolean equations for parallel–serial connection, see, e.g., [2.1], [2.13]. However, this procedure has not proved to be successful because of the continuous and gradual nature of faults and symptoms. For the diagnosis of technical processes approximate reasoning is more appropriate. A recent survey and learning methods for rule-based diagnosis is given in [2.16] and [2.15].

2.7 Fault detection and diagnosis in closed loop

The main goals for using automatic control loops are precise following of reference variables (setpoints), a faster response than in open loop, compensation of all kind of external disturbances on the controlled variable, stabilization of unstable processes, reduction of the influence of process parameter changes with regard to the static and dynamic behavior, partial compensation of actuator and process nonlinearities, and, of course, replacement of manual control by humans. The performance of a SISO control loop with regard to the control error (deviation)

$$e(k) = w(k) - y(k) \tag{2.7.1}$$

i.e. the deviation of the controlled variable $y(k)$ from the reference variable $w(k)$ depends on many facts, compare Figure 2.19, like:

- external disturbance $w(k), u_v(k), v_i(k)$
- structure and parameters of the controller G_c and controller faults f_c
- changes of the structure and parameters of the process G_p and process faults f_p

- changes and faults of actuator G_a and f_a
- faults f_s in the sensor G_s and measurement noise n_s.

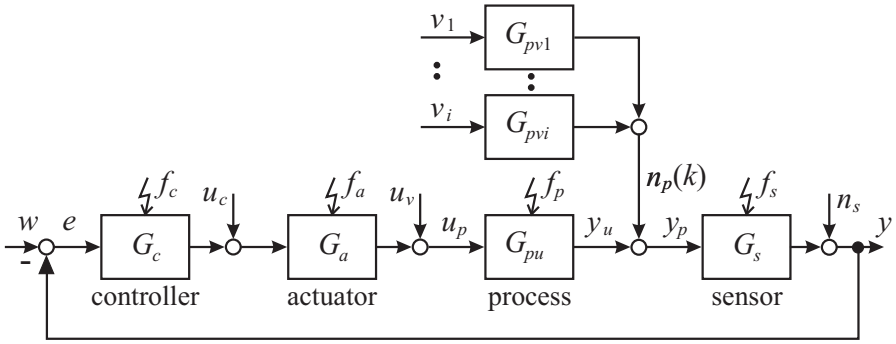


Fig. 2.19. Control loop with variables and fault influences

y	controlled variable	w	reference variable
u_p	manipulated variable	e	control deviation
v_i	process disturbances	n_s	measurement noise
u_v	process input disturbances	$f_{c,a,p,s}$	faults of the controller,
n_p	sum of process disturbances		actuator, process and
y_p	process output to be controlled		sensor

Hence, many changes and faults influence the performance of closed loops. Usually, only the control deviation e and the control variable y are monitored.

Small faults in the actuator and process, be they additive or multiplicative, will usually be compensated by the feedback controller (with integral action) and they will not be detectable by considering $e(k)$ and $y(k)$ only, as long as the control deviation turns back to approximately zero. Also small sensor offset faults will not be detected. The controller will just make the wrong sensor signal equal to the reference variable. Only by a redundant sensor or other redundant information for the controlled variable, can the offset fault usually be detected.

As shown in [2.37], Chapter 12, several *larger faults* have a similar effect on the considered changes of closed-loop behavior such that it is not easy to differentiate them. In addition, some of the behavior is also observed after external disturbances under normal operation.

A first possibility for fault detection in closed loops is to analyze measurable signals like the controlled variable $y(t)$, the manipulated variable $u(k)$, the reference variable $w(k)$ and the control deviation $e(k)$. This is also known as *performance monitoring* of closed loops, see, e.g. [2.37], Chapter 12. For example, the variances of these variables, steady-state deviations, large overshoots or frequency spectra can be monitored. However, it is very difficult to find the reasons for observed changes.

Process-model-based detection methods allow a deeper insight, as they relate the manipulated variable $u(k)$ to the controlled variable $y(k)$. The application of *parameter estimation* in closed loops requires consideration of the identifiability

conditions. If no external measurable perturbations can be used, because the closed loop operates with constant reference variable $w(k) = const$ and only compensates for disturbances, special higher-order controller structures are required, [2.40]. With measurable external perturbations, as for servo systems with continuously changing reference variable, parameter estimation methods can be directly applied. Also *parity equations* for the process are directly applicable if the (fixed) process model agrees well with the real process. Especially by combining several detection methods a large portion of faults in the components of a closed loop (actuators, process, sensor, controller) can be detected and isolated [2.37].

2.8 Data flow structure for supervision (condition monitoring)

For larger plants the computer software systems written for supervision or condition monitoring and diagnosis of machines and other processes should have an open standardized software architecture to ease the exchange of data. This has to be seen in the efforts to integrate the components and processes with regard to a plant asset management. Therefore, international standards are prepared, like ISO 13374 on "condition monitoring and diagnostics of machines," [2.41] with regard to data processing, communication, and presentation.

An *open condition monitoring software architecture* then consists of a standardized structure for data processing with an information model, a data model and a data library. The *information model* (data flow scheme) describes the primary data objects with their properties (attributes) in the form of a scheme which is independent of the physical data. The information model can, e.g. be implemented with the Unified Modeling Language (UML), containing standardized class diagrams for information modeling.

The *data model* is based on the information model and provides the exact representation of data elements. Here, the integration of many sources of machinery or process information takes place, like process site, asset nameplate data for rated quantities, measurement locations, signal processing methods, alarms, date and time. An Extensible Markup Language (XML) is a recommended definition language.

A *reference data library* then stores the data from the data model, using database-unique entries. This library specifies all code tables, asset types, event codes, health codes, failure codes, root cause codes and engineering unit codes.

The software structure for *data processing* is recommended to follow Figure 2.20. The *data acquisition* digitizes, e.g. analog sensor signals and manual inputs and delivers digitized data with time stamps and data quality (good, bad, unknown). The digital data are then processed in the *data manipulation block* providing specific features with time stamp and data quality. The signal processing consists of algorithmic computations, filtering, windowing, spectral analysis and feature extraction.

A *state detection block* then categorizes the features with regard to normal or abnormal, exceeding of limits, severity of threshold boundary deviation, degree of abnormality and statistical analysis. The *health assessment* determines the current

state of health and potential failures with associated diagnosis, calculates the current risk priority number and generates recommendations.

The *prognostic assessment* projects the future health state with prognostic models, future operational usage, failure rates, probability measures and explanations.

Finally, the *advisory generation* integrates all information and provides optimal recommended actions and alternatives, maintenance, modification of operations, capability forecast, strategic recommendations.

All the blocks have a time stamp and need multiple interactions, require configuration information and the access to external information sources with previous maintenance, historic operational data and should be accessible from displays for deeper analysis.

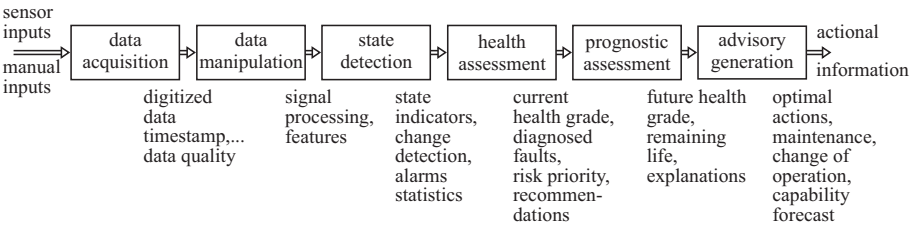


Fig. 2.20. Data processing functionality, scheme for condition monitoring and diagnosis, [2.42]

Concluding remarks

The summary of some basic fault-detection and diagnosis methods presented in this chapter was limited to linear processes mainly. Some of the methods can also be directly applied to nonlinear processes, e.g., signal analysis, parity equations and parameter estimations. However, all the methods have to be adapted to the real processes. In this sense the basic methods should be considered as "tools", which have to be combined properly in order to meet the practical requirements for real faults of real processes. The *development of fault detection and diagnosis methods* is recommended to follow the schedule depicted in Figure 2.21. First, the *requirements* of the final results should be stated, where especially a list of all faults to be detected is defined and fault-symptom trees are sketched, supported by an FMEA. A *process analysis* has to follow, stating the available measurements and operating conditions. Then, possible *fault-detection methods* can be selected. *Simulations* with process and signal models allow first tests. Important are, of course, *experiments with the real process*, real-time computing and handling of real measured data. The results give hints for adjustments and improvements of detection methods in the sense of *feedback* and iterative scaling. A *final realization* of software and computer hardware usually requires practical tests with similar or other processes, solving of robustness issues and field tests.

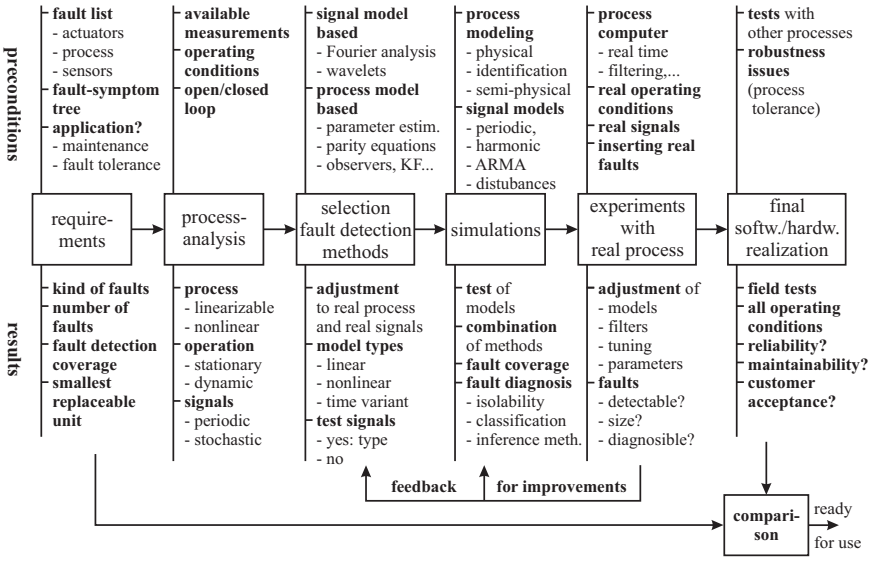


Fig. 2.21. Development stages for fault detection and diagnosis

Most of the described fault-detection and diagnosis methods, see [2.37], were developed theoretically, investigated by simulations and then tested experimentally on real processes, either in laboratories on test rigs or with pilot processes and different machinery as described in the following chapters. In some cases the faults are added artificially, like for sensor offsets, but in many cases real faults are introduced, if the faults did not damage the processes.

Part II

Drives and Actuators

3

Fault diagnosis of electrical drives

Electrical drives are basic components in a multitude of devices, processes, machinery and vehicles, and in the large areas of mechanical power and process engineering, manufacturing, transportation and precision mechanical devices. Their power ranges from a few mW to hundreds MW.

The most important types of electrical motors can be divided into:

(i) DC motors
- series-wound motors
- shunt-wound motors
- permanent-field motors

(ii) Three-phase AC motors
- induction motors (asynchronous motors)
- synchronous motors

(iii) Single-phase AC motors
- commutator motors (universal motors)
- squirrel-cage motors.

Table 3.1 gives an overview of some basic types, illustrating torque characteristics and corresponding control inputs. As static and dynamic models of the various electrical motors are required for model-based fault detection, the reader is referred to well-known basic books on electrical drives such as [3.3], [3.13], [3.18], [3.19], [3.20].

In the following, some case studies are described for DC motors with brushes and for AC motors. Further types of electrical motors will be considered in Chapter 4 for electrical actuators.

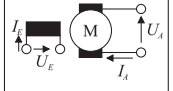
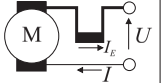
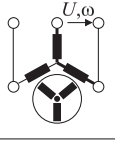
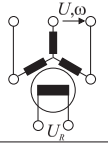
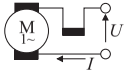
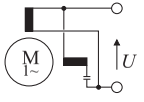
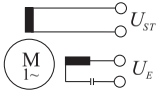
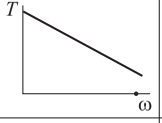
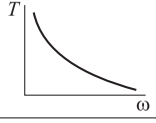
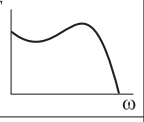
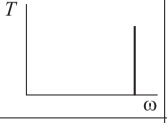
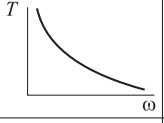
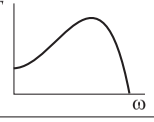
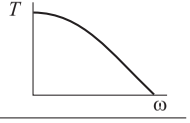
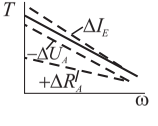
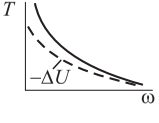
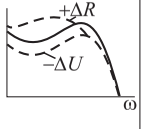
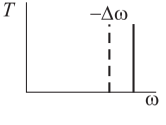
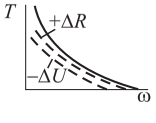
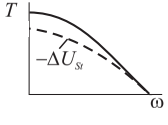
3.1 Direct-current motor (DC)

3.1.1 Structure and models of a DC motor

A permanently excited DC motor with a rated power of $P = 550\,\mathrm{W}$ at rated speed $n = 2500\,\mathrm{rpm}$ is considered, [3.6]. This DC motor has a two-pair brush commuta-

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_3, © Springer-Verlag Berlin Heidelberg 2011

Table 3.1. General survey of electrical motors with small power, [3.9]

motor		DC shunt-wound motor	DC series-wound motor	three-phase asynchronous motor	three-phase synchronous motor (DC excitation of rotor)	single-phase (universal motor)	single-phase asynchronous motor with condensator	single-phase asynchronous motor (Ferraris motor)
circuit diagram								
torque-speed characteristics								
torque characteristics for manipulation	different manipul. variables:							
	normal:							
	manipulated variables	ΔU_A armature voltage ΔI_E excitation current ΔR_A armature resistance	ΔU voltage	ΔU voltage $\Delta\omega$ frequency ΔR rotor resistance	$\Delta\omega$ frequency	ΔU voltage ΔR armature resistance		ΔU_{St} manipulated voltage

tion, two pole pairs, and an analog tachometer for speed measurement; it operates against a hysteresis brake as load, see Figure 3.1. The measured signals are the armature voltage U_A, the armature current I_A and the speed ω. A servo amplifier with pulse-width-modulated armature voltage as output and speed and armature current as feedback allows a cascaded speed control system. The three measured signals first pass analog anti-aliasing filters and are processed by a digital signal processor (TXP 32 CP, 32-bit fpt, 50 MHz) and an Intel Pentium host PC. Also the hysteresis brake is controlled by a pulse-width servo amplifier. Usually such DC motors can be described by linear dynamic models.

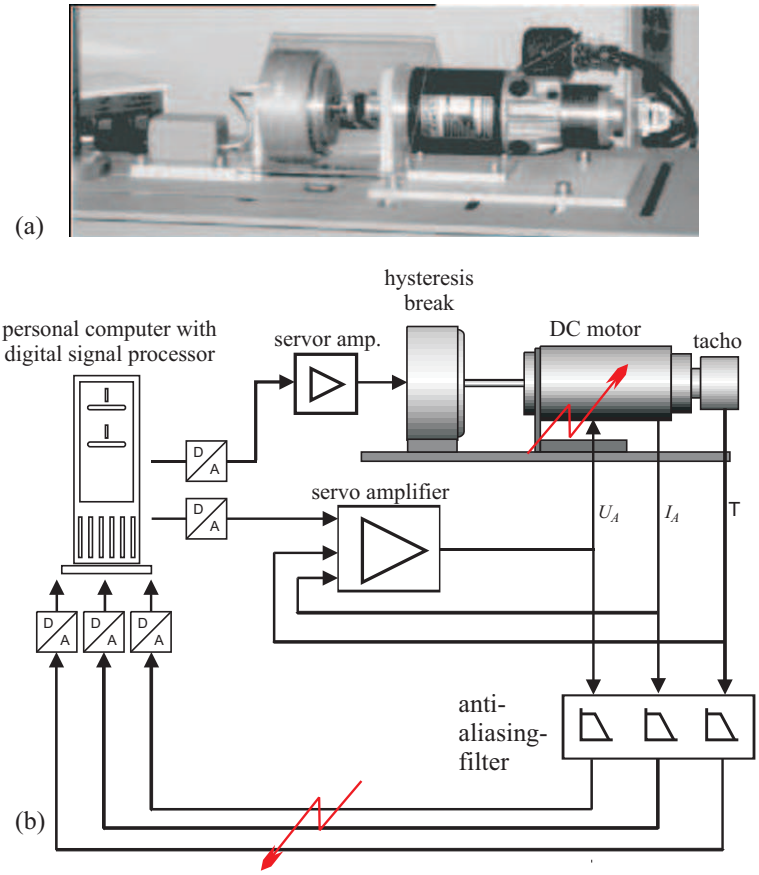


Fig. 3.1. DC motor test bench with hysteresis brake: a) test bench; b) scheme of equipment

However, experiments have shown that this model with constant parameters does not match the process in the whole operational range. Therefore, two nonlinearities are included so that the model fits the process better. The resulting first-order differential equations are:

$$L_A\ \dot{I}_A(t) = -R_A\ I_A(t) - \Psi\ \omega(t) - K_B|\omega(t)|\ I_A(t) + U_A^*(t) \tag{3.1.1}$$

$$J\ \dot{\omega} = \Psi\ I_A(t) - M_{F1}\ \omega(t) - M_{F0}\ \text{sign}\ (\omega(t)) - M_L(t) \tag{3.1.2}$$

Figure 3.2 depicts the resulting signal flow diagram. The term $K_B|\omega(t)|I_A(t)$ compensates for the voltage drop at the brushes in combination with a pulse-width-modulated power supply. The friction is included by a viscous- and a dry-friction term $M_{F1}\omega$ and $M_{F0}\text{sign}(\omega)$, see also [3.9]. The parameters are identified by least-squares estimation in the continuous-time domain, [3.6]. Table 3.2 gives the nominal values. Most of them ($R_A, \Psi, K_B, M_{F1}, M_{F0}$) influence the process gain, and the other two (L_A, J) the time constants. The signals U_A^*, I_A and ω are measured with a sampling frequency of 5 kHz, and state-variable filtered by a fourth-order low-pass filter with Butterworth characteristic and a cut-off frequency of 250 Hz.

Table 3.2. Data for the DC motor

armature resistance	$R_A = 1.52\ \Omega$
armature inductance	$L_A = 6.82 \cdot 10^{-3}\ \Omega\ \text{s}$
magnetic flux	$\Psi = 0.33\ \text{V s}$
voltage drop factor	$K_B = 2.21 \cdot 10^{-3}\ \text{V s / A}$
inertia constant	$J = 1.92 \cdot 10^{-3}\ \text{kg m}^2$
viscous friction	$M_{F1} = 0.36 \cdot 10^{-3}\ \text{Nm s}$
dry friction	$M_{F0} = 0.11\ \text{Nm}$

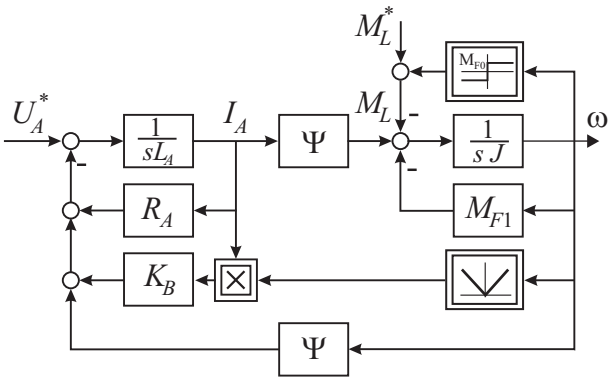


Fig. 3.2. Signal flow diagram of the considered DC motor

3.1.2 Fault detection with parity equations

For the detection and isolation of sensor (output) and actuator (input) faults a set of structured parity equations with state-space models according to Section 2.5 is applied.

As the differential equations (3.1.1) and (3.1.2) are nonlinear, the design procedure for a linear parity space cannot be applied directly. But defining $U_A^* - K_B|\omega(t)|I_A$ as voltage input U_A and as load input $M_L = M_{F0}\,\mathrm{sign}\,\omega$ leads to a linear description. The linear state-space representation then becomes

$$\begin{aligned}\dot{\mathbf{x}} &= \begin{bmatrix} \dot{I_A} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} -\frac{R_A}{L_A} & -\frac{\Psi}{L_A} \\ \frac{\Psi}{J} & -\frac{M_F}{J} \end{bmatrix} \begin{bmatrix} I_A \\ \omega \end{bmatrix} + \begin{bmatrix} \frac{1}{L_A} & 0 \\ 0 & -\frac{1}{J} \end{bmatrix} \begin{bmatrix} U_A \\ M_L \end{bmatrix} \\ \mathbf{y} &= \begin{bmatrix} I_A \\ \omega \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x} \end{aligned} \tag{3.1.3}$$

A corresponding signal flow diagram is depicted in Figure 3.2.

An observability test reveals that both outputs (I_A and ω) can also observe each other. This is a precondition for a parity space of full order (here: 2). Then, $\mathbf{W}$, see Table 2.4 and [3.10] Equation (10.52), is chosen such that a set of *structured residuals* is obtained, where residual $r_1(t)$ is independent of $M_L(t)$, $r_2(t)$ of $U_A(t)$, $r_3(t)$ of $\omega(t)$ and $r_4(t)$ of $I_A(t)$, see also [3.6], [3.16], [3.4]:

$$\mathbf{W} = \begin{bmatrix} R_A & \Psi & L_A & 0 & 0 & 0 \\ -\Psi & M_{F1} & 0 & J & 0 & \\ \alpha & 0 & \beta & 0 & JL_A & 0 \\ 0 & \alpha & 0 & \beta & o & JL_A \end{bmatrix} \tag{3.1.4}$$

with $\alpha = \Psi^2 + R_A\, M_{F1}$;
$\beta = L_A\, M_{F1} + J\, R_A$.
The residuals, using three measured signals, then follow as:

$$\begin{aligned} r_1(t) &= L_A\, \dot{I}_A(t) + R_A\, I_A(t) + \Psi\, \omega(t) - U_A(t) \\ r_2(t) &= J\, \dot{\omega}(t) - \Psi\, I_A(t) + M_{F1}\, \omega(t) + M_L(t) \\ r_3(t) &= J\, L_A\, \ddot{I}_A(t) + (L_A\, M_{F1} + J\, R_A)\, \dot{I}_A(t) \\ &\quad + (\Psi^2 + R_A M_{F1}) I_A(t) - J\, \dot{U}_A(t) - M_{F1}\, U_A(t) - \Psi\, M_L(t) \\ r_4(t) &= J\, L_A\, \dddot{\omega}(t) + (L_A\, M_{F1} + J\, R_A)\, \ddot{\omega}(t) + (\Psi^2 + R_A\, M_{F1})\, \omega(t) \\ &\quad - \Psi\, U_A(t) + L_A\, \dot{M}_L(t) + R_A\, M_L(t) \end{aligned} \tag{3.1.5}$$

The same residual equations can be also obtained via transfer functions as described in Example 10.3 in [3.10]. If an additive fault of the measured signals and of M_L occurs, all residuals except the decoupled one are deflected. The scheme of the structured residuals is not touched by the compensation for the nonlinear voltage drop of the brushes, as its magnitude is small enough. Two parameters R_A and M_{F1}, however, depend on the present motor temperature. The behavior of R_A and its effect on residual r_1 is depicted in Figure 3.3. Therefore, the use of adaptive parity equations improves the residual performance, see [3.6] and [3.10].

The residuals are now examined with regard to their sensitivity to additive and parametric faults. As r_1 and r_2 comprise all parameters and all signals, it is sufficient to consider only these two, although r_3 or r_4 can also be taken. From (3.1.5) it yields

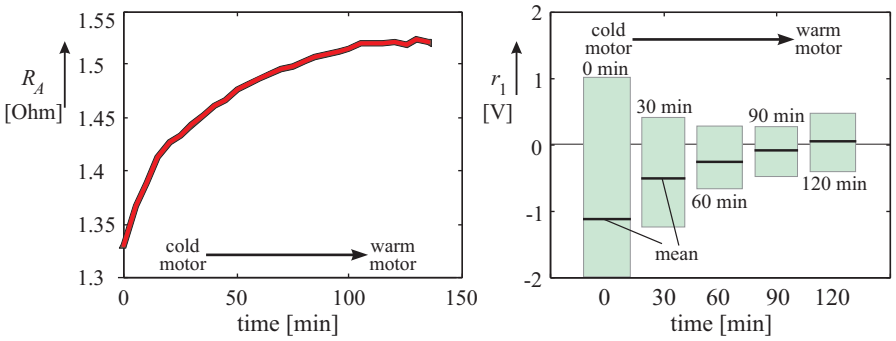


Fig. 3.3. Influence of the motor temperature on resistance R_A and residual r_1

$$\begin{aligned} r_1(t) &= \Delta\, L_A\, \dot{I}_A(t) + \Delta\, R_A\, I_A(t) + \Delta\, \Psi\, \omega(t) \\ &\quad + L_A\, \Delta\, \dot{I}_A(t) + R_A\, \Delta\, I_A(t) + \Psi\, \Delta\, \omega(t) - \Delta\, U_A(t) \\ r_2(t) &= +\Delta\, J\, \dot{\omega}(t) - \Delta\, \Psi\, I_A(t) + \Delta\, M_{F1}\, \omega(t) \\ &\quad + J\, \Delta\, \dot{\omega}(t) - \Psi\, \Delta\, I_A(t) + M_{F1}\, \Delta\, \omega(t) + \Delta\, M_L(t) \end{aligned} \tag{3.1.6}$$

In the presence of residual noise, e.g. of r_1 with a magnitude of about 1 V and an armature current of 3 A, a resistance change must be at least 0.3 Ω in order to deflect the residual significantly. Therefore, the two linear parameters R_A and M_{F1} are selected to be tracked according to a single parameter estimation together with parity equations, as described in [3.10], Section 10.5. The forgetting factor is chosen as $\lambda = 0.99$.

3.1.3 Fault detection with parameter estimation

The parameter estimation is based on the two differential equations (3.1.1) and (3.1.2) in the simplified form

$$\dot{I}_A(t) = -\hat{\theta}_1\, I_A(t) - \hat{\theta}_2\, \omega(t) + \hat{\theta}_3\, U_A(t) \tag{3.1.7}$$

$$\dot{\omega}(t) = \hat{\theta}_4\, I_A(t) - \hat{\theta}_5\, \omega(t) - \hat{\theta}_6\, M_L(t) \tag{3.1.8}$$

with the process coefficients

$$R_A = \frac{\hat{\theta}_1}{\hat{\theta}_3};\; L_A = \frac{1}{\hat{\theta}_3};\; \Psi = \frac{\hat{\theta}_2}{\hat{\theta}_3} \text{ and } \Psi = \frac{\hat{\theta}_4}{\hat{\theta}_6};\; J = \frac{1}{\hat{\theta}_6};\; M_{F1} = \frac{\hat{\theta}_5}{\hat{\theta}_6} \tag{3.1.9}$$

Applying the recursive parameter estimation method DSFI (discrete square-root filtering in information form), [3.11], with forgetting factor $\lambda = 0.99$ yields the parameters $\hat{\theta}_i$ by using three measured signals. Then all process coefficients can be calculated with (3.1.9). Experimental results with idle running ($M_L = 0$) resulted in standard deviations of the process coefficients in the range of $2\% < \sigma_\theta < 6.5\%$, [3.6].

3.1.4 Experimental results for fault detection (SELECT)

Based on many test runs, five different faults are now selected to show the detection of additive and multiplicative faults with parity equations and recursive parameter estimation, [3.5]. The time histories depict the arising faults at $t = 0.5\,\mathrm{s}$. The faults are step changes and were artificially produced. Figure 3.4 shows the parameter estimates and the residuals of parity equations. The residuals are normalized by division through their thresholds. Therefore, exceeding of 1 or -1 indicates the detection of a fault. In the cases a) to d) and f) the DC motor is excited by a pseudo random binary signal (PRBS) of the armature voltage U_A which is a requirement for dynamic parameter estimation, as shown in Figure 3.4f). In case e) the input is constant. The results can be summarized as:

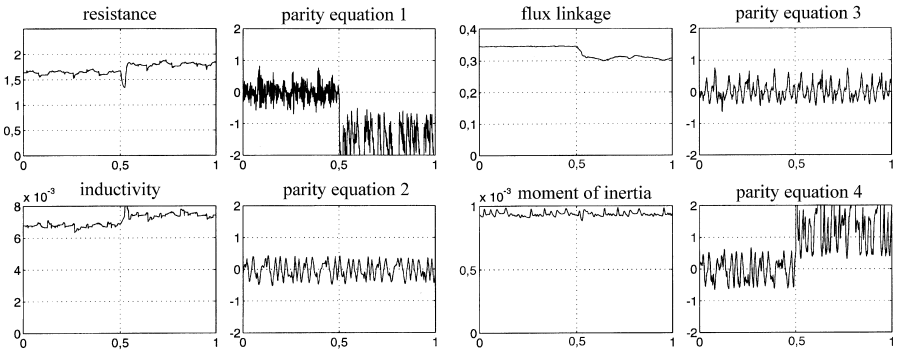
a) A sensor-gain fault of the voltage sensor U_A leads as expected to a change of residual 1 (and 3, 4) but not of residual 2, which is independent of U_A. The parameter estimates show (incorrect) changes for R_A, L_A and Ψ, because the gain of the voltage sensor is not modeled
b) An offset fault in the speed sensor ω leads to a change of the residuals r_4, r_1 and r_2, but r_3 remains uneffected, because it is independent of ω. The parameter estimate of Ψ shows an (incorrect) change
c) A multiplicative change of the armature resistance R_A yields a corresponding change of the parameter estimate $\hat{R}_A$. However, the residuals increase their variance drastically and exceed their thresholds
d) A change of the ratio of inertia is correctly given by the parameter estimate $\hat{J}$. But all residuals, except r_1, exceed their thresholds by increasing their variance
e) The same fault in R_A as in c) is introduced, but the input U_A is kept constant. The parameter estimate $\hat{R}_A$ does not converge to a constant value and the parity equation residuals r_1 and r_4 change their mean, however, with large variance
f) A brush fault leads to an increase of R_A and L_A but not of Ψ. The residuals show an increase of the variance.

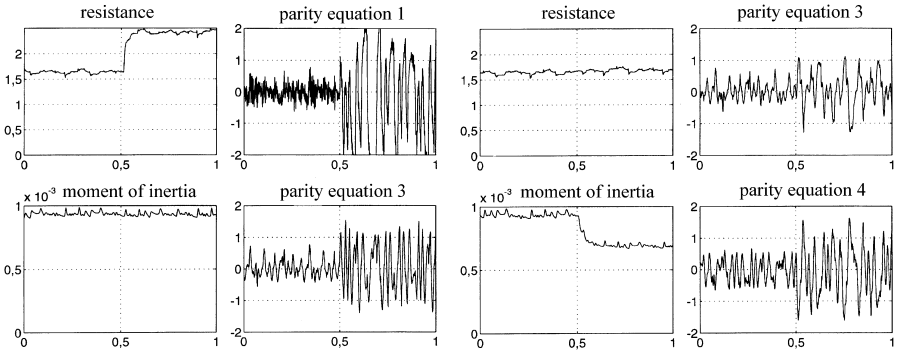
Table 3.3 summarizes the effects of some investigated faults on the parameter estimates and parity residuals.

These investigations have shown:

1) *Additive faults* like the offsets of sensors are well detected by the parity equations. They react fast and do not need an input excitation for a part of the faults. However, they have a relatively large variance, especially if the model parameters do not fit well to the process
2) *Multiplicative faults* are well detected by parameter estimation, also for small faults. Because of the inherent regression method the reactions are slower but smoothed. But they require an input excitation for dynamic process models.

Therefore, it is recommended to combine both methods, as shown in [3.10], Section 14.3. The parity equations are used to detect changes somewhere in the process and if the fault detection result is unclear a parameter estimation is started, eventually by a dynamic test signal for some seconds. If the motor operates dynamically




resistance
parity equation 1
moment of inertia
parity equation 3
(c) change of armature resistance (ΔR_A : +50%)
resistance
parity equation 3
moment of inertia
parity equation 4
(d) change of moment of interia (ΔJ : -30%)

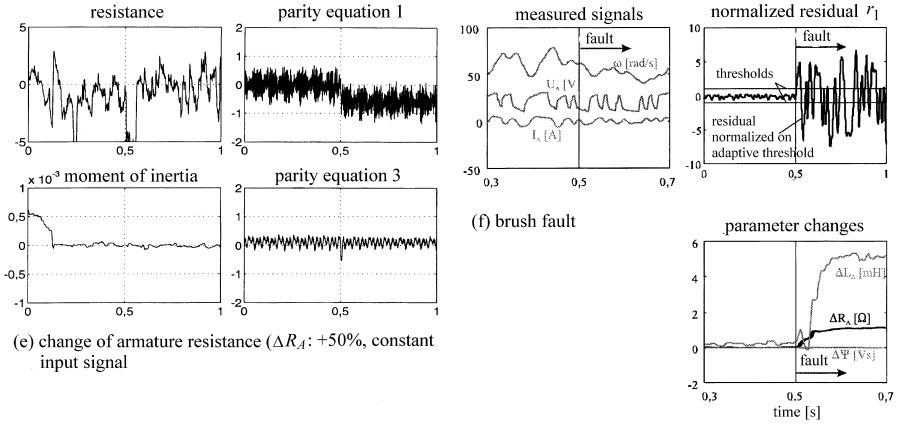


Fig. 3.4. Time histories of signals, residuals of parity equations and parameter estimation at fault occurrence

- parameter estimates: R_A resistance, L_A inductivity, Ψ flux linkage, J_A moment of inertia
- parity equations: r_1, r_2, r_3, r_4

Table 3.3. Fault-symptom table for the fault detection of a DC motor with dynamic input excitation $U_A(t)$ in the form of a PRBS. + positive deflection; ++ strong positive deflection; 0 no deflection; − negative deflection; − − strong negative deflection; ± increased variance

faults			symptoms								
			parameter estimation					parity equations			
			R_A	L_A	Ψ	J	M_{F1}	r_1	r_2	r_3	r_4
parametric faults	armature resistance	ΔR_A	++	0	0	0	0	±	0	±	±
	brush fault		++	+	0	0	0	±	0	±	±
	change of inertia	ΔJ	0	0	0	++	0	0	±	±	±
	change of friction	ΔM_{F1}	0	0	0	0	++	0	±	±	±
	voltage sensor gain fault	ΔU_A	±	±	±	0	0	−	0	−	−
additive faults	speed sensor offset fault	$\Delta\omega$	0	0	−	0	0	+	+	0	+
	current sensor offset fault	ΔI	±	±	±	0	0	+	−	+	0

anyhow (as for servo systems and actuators) then the parameter estimation can be applied continuously, but with a supervision scheme, see [3.11].

[3.6] has shown that a considerable improvement can be obtained by continuously estimating the armature resistance with a single parameter estimation using parity equations in order to reach the temperature dependent resistance parameters, [3.7]. Furthermore, adaptive thresholds are recommended, to compensate for model uncertainties, see Section 2.4.4.

3.1.5 Experimental results for fault diagnosis with a learning fault-symptom tree

The model-based fault-detection system with parity equations and parameter estimation is now the basis for a fault-diagnosis procedure. As described in Section 2.6 the method for fault diagnosis can be divided in classification and inferencing. A first simple classification is the use of fault-symptom tables and pattern recognition as in Table 3.3. Also decision trees belong to the class of classification methods. However, a combination with a neuro-fuzzy structure gives them a learning behavior of fuzzy if-then rules with AND operators, forming an adaptive inference method, called SELECT, [3.4]. This is applied in the following to the DC motor test bench.

a) The symptoms used

To diagnose the faults, altogether 22 symptoms are created:

- Windowed sums of the absolute values of the three measured signals U_A^*, I_A, ω
- Mean values and standard deviations of four residuals: $\bar{r}_1, \ldots, \bar{r}_4$ and $\bar{\sigma}_{r1}, \ldots, \bar{\sigma}_{r4}$

- Eight parameter estimates. Symptoms are the deviations of the current values – results of the estimation – from the nominal ones. They are normalized to the nominal values. For the rotor resistance R_A this is $\Delta R_{A1} = \frac{(R_{A,nom.} - R_{A,est.})}{R_{A,nom.}}$. The index 1 denotes that the estimation was carried out using the first parity equation. Similarly, ΔR_{A4}, ΔL_{A1}, ΔL_{A4}, ΔJ_2, ΔJ_3, ΔM_{F12}, and ΔM_{F13} are computed
- Additionally, two symptoms judge the quality of the estimation. They describe the variance of an estimated parameter during a recursive estimation. This variance can give a good indication whether the structure of the estimation equation is valid. A structural change of the system will result in a bad estimation result where the recursively estimated parameters fluctuate significantly. Two parameter estimations were chosen: Ψ and M_{F1}. Their estimation variances are denoted by $\sigma_{est.,\Psi}$ and $\sigma_{est.,MF1}$.

The symptoms serve to differentiate between 14 fault situations that can artificially be introduced on the test rig.

The DC motor diagnosis was performed by learning a SELECT tree from experimentally gained fault data. For the fault cases, typically 10–50 test-cycle measurements for a parameter estimation were performed. The residuals were computed from the test runs. That way, each test run results in one data point in the symptom space. The membership functions were created with the degressive fuzzy-c-means method. To utilize a maximum of transparency and create a highly interpretable system, prior knowledge was used to structure the diagnosis system.

b) Incorporation of structural knowledge

In most applications, a certain amount of knowledge about the symptom behavior is present. Even if exact values for thresholds etc. are not known, there usually is some insight into the process like physical understanding of similar faults or similar effects of faults on certain symptoms. For the DC motor, this could be as simple as to use the windowed sums of the signals in order to to detect a broken sensor cable. This information is quite obvious, but its benefits are sometimes neglected, if a diagnosis system is designed with the aim to be solely learned from measured data. Hence, the task could be simpler if the designer used this information from the beginning.

Furthermore, the selection of the symptoms for the diagnosis becomes a matter of *robustness*. Some symptoms are affected by faults for which they are not an appropriate indicator. In an experimental environment, it is virtually impossible to gather enough measurements to adequately reflect every influence. Especially changes in the environmental conditions and long-term changes due to wear are hardly captured in a limited time frame. This leads to diagnosis systems that work well under the experimental conditions but fail otherwise. The diagnosis of a fault should therefore be based on the *appropriate subset* of all available symptoms. Only the relevant ones should be selected.

Often, different faults can be categorized into larger groups if their effects on the process are similar. It is then advantageous to find a classification system for

the larger groups first and later separate within them. This leads to the concept of a *hierarchical* diagnosis system.

Overall, it is proposed to use prior knowledge to *structure* the diagnosis system. The designer builds groups of faults and identifies the corresponding relevant symptoms to first differentiate between and later within them. The exact decisions can be found automatically if enough measured data is available.

If the set of all different fault situations F_i is denoted by

$$\mathcal{F} = \{F_1, F_2, \ldots F_r\} \tag{3.1.10}$$

and the available symptoms given by

$$\mathcal{S} = \{s_1, s_2, \ldots s_t\} \tag{3.1.11}$$

one can form meta-classes $\mathcal{C}_i, i = 1 \ldots m$ with

$$\mathcal{F} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \ldots \cup \mathcal{C}_m \tag{3.1.12}$$

In the DC motor diagnosis, for instance, such a meta-class is given by all faults on the mechanics of the motor. Such a hierarchy based on meta-classes requires at least $q = m + r$ decisions d_j, $j = 1 \ldots q$ assumed that no $\mathcal{C}_i$ is a single-element set. Each d_j is based on a subset $\mathcal{S}_{dj} \in \mathcal{S}$. The SELECT approach will then produce a system with p parameters where p is given by

$$p = \sum_{j=1}^{q} \text{card}\,(\mathcal{S}_{dj}) \tag{3.1.13}$$

which is typically much less than a parallel network structure would result in (cardinalities are the number of relevant sets). The usually larger number of parameters in parallel network configurations can lead to slower convergence and ill-conditioned optimization problems.

In addition to the structural knowledge, one can incorporate more detailed knowledge into the individual rules if desired.

c) Results with SELECT method

A total of 14 different fault situations are applied on the DC motor test bench:

- Change of rotor inductance or resistance F_{RA}, F_{LA}
- Broken rotor wiring (F_W)
- Failure of one the four brushes (F_B)
- Increased friction in the bearings (F_F)
- Offset on voltage, current or speed sensor signal ($F_{O,UA}$, $F_{O,IA}$, $F_{O,\omega}$)
- Gain change of voltage, current or speed sensor signal ($F_{G,UA}$, $F_{G,IA}$, $F_{G,\omega}$)
- Complete voltage, current or speed sensor failure (F_{UA}, F_{IA}, F_{ω}).

Repeated experiments with different faults were performed using a test cycle. The symptoms described in a) were computed for each of the experiments. Overall, the training set for the approach consisted of data from 140 experiments.

Figure 3.5 shows the resulting structure for the DC motor diagnosis. Details have been omitted to visualize the concept only. Each block comprises a meta-class $\mathcal{C}_1$ of faults. Every branching of the tree is connected to a decision d_j learned with the SELECT approach, i.e. it contains a fuzzy rule. In each meta-class, a classification tree decides which individual fault has occurred based on a subset $\mathcal{S}_i$ of the symptoms.

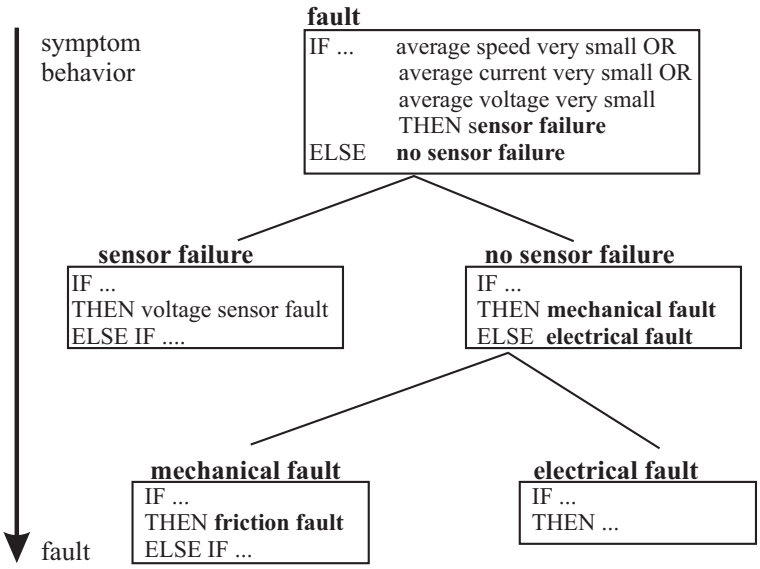


Fig. 3.5. Hierarchical fault-diagnosis system. Each block comprises a fuzzy classification tree

The hierarchical decision tree proved to be highly suitable for the diagnosis. It achieved a 98% classification rate in a cross-validation scheme.

The groups of faults have been selected following basic understanding of the DC motor supervision concept. Firstly, the three total sensor breakdowns are different from other faults due to their strong effects on all symptoms. They form the first meta-class $\mathcal{C}_1$ and can be easily differentiated by the three windowed sums of the signals. These three symptoms accordingly form the set $\mathcal{S}_1$.

Since the motor can be understood as a combination of an electrical and a mechanical component, faults on these two parts were again treated separately, creating two more meta-classes, $\mathcal{C}_2$ and $\mathcal{C}_3$. Accordingly, the appropriate subsets of symptoms $\mathcal{S}_2$ and $\mathcal{S}_3$ for the diagnosis were selected. Basically, $\mathcal{S}_2$ and $\mathcal{S}_3$ consist of the residuals and parameter deviations connected to the corresponding meta-class. The diagnosis of electrical faults, for instance, is not based on parameter estimates of the mechanical parameters. Although some electrical faults may have an influence on the estimates of the mechanical parameters, this influence should not be

used as the estimates are misleading and not reliable. Hence, $\mathcal{S}_2$ does not contain $\Delta J_2, \Delta J_3, \Delta M_{F12}$ or ΔM_{F13}.

To give an example of the SELECT approach, the rules for the distinction of the electrical faults are given below:

$$\begin{aligned}
&\text{IF } \bar{r}_1 \text{ is small AND } \Delta L_{A4} \text{ is strongly negative THEN Fault } F_{LA}\\
&\text{ELSE IF } \bar{r}_1 \text{ is small AND } \bar{\sigma}_{r4} \text{ is medium THEN Fault } F_{RA}\\
&\text{ELSE IF } \bar{r}_1 \text{ is small AND } \bar{\sigma}_{r4} \text{ is large THEN Fault } F_B\\
&\text{ELSE IF } \bar{r}_2 \text{ is not small THEN Fault } F_{0,IA}\\
&\text{ELSE IF } \bar{r}_1 \text{ is small THEN Fault } F_{G,IA}\\
&\text{ELSE IF } \bar{r}_1 \text{ is large AND } \sigma_{est.,\Psi} \text{ is not small THEN Fault } F_{0,UA}\\
&\text{ELSE Fault } F_{G,UA}
\end{aligned} \tag{3.1.14}$$

The relevance indices of the rule premises are not listed here. They also play a role for the exact decision boundaries.

Nevertheless, it is possible to analyze and understand parts of these rules. Clearly, the rules reveal the discriminatory power of the first residual, since it was used very often. Other rule premises are also understandable. The change of the rotor inductance is indicated by a strongly negative estimation of this change magnitude. Compare this rule to Figure 3.6a). It shows the values ΔL_{A4} for the electrical faults from the training set. Clearly, the fault F_{LA} makes a distinct difference. Hence, it makes sense to use ΔL_{A4} to distinguish the fault from the others. The corresponding membership functions are shown in Figure 3.6b). It must be noted that the experimental setup allowed only a fixed deviation of the inductance by –50% as a fault. That can be seen in the estimation result. If, however, also positive changes are to be diagnosed, one is able to enhance the rule manually. For instance, one could use

$$\text{IF } \bar{r}_1 \text{ is small AND } \Delta L_{A4} \text{ is not small THEN Fault } F_{LA} \tag{3.1.15}$$

The corresponding membership functions for ΔL_{A4} would also have to be adapted accordingly to allow processing of positive values of ΔL_{A4}.

Another interesting observation is the use of $\sigma_{est.,\Psi}$ in the sixth rule of (3.1.14) to distinguish offset from gain faults of the voltage sensor. This can be explained by the fact that an offset term in the estimation equation given by an offset fault will change the *structure* of the estimation equation, while a gain will only effect parameters. Hence, the normal estimation equation will still be valid in the case of gain faults, but indicate a problem by a large $\sigma_{est.,\Psi}$ for offset faults.

The system performed well on new experiments, showing the increased robustness through the incorporation of very simple knowledge. Additionally, the system has a higher degree of transparency facilitating an adaptation to other motors. The diagnostic rules can be extracted and are largely understandable.

d) Relation to fault trees

The resulting hierarchical classifier can also be interpreted as a set of fuzzy fault trees. If one *reverses* the order of the structure and traces the decisions leading to a

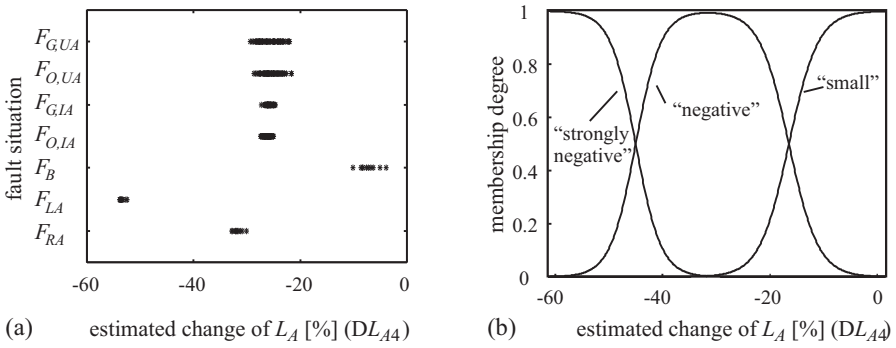


Fig. 3.6. Estimated rotor inductance computed from the fourth parity residual. Apparently, most faults influence the result, however, the faulty inductance can most easily be detected due to its strong influence: a) estimation results; b) resulting membership functions

particular fault back through the tree, it is possible to explicitly draw a fault tree for each individual fault. Figure 3.7 shows one fault situation (increased friction in the motor) as an example. The intermediate steps like "mechanical fault" from Figure 3.7 become *events* of the fault tree.

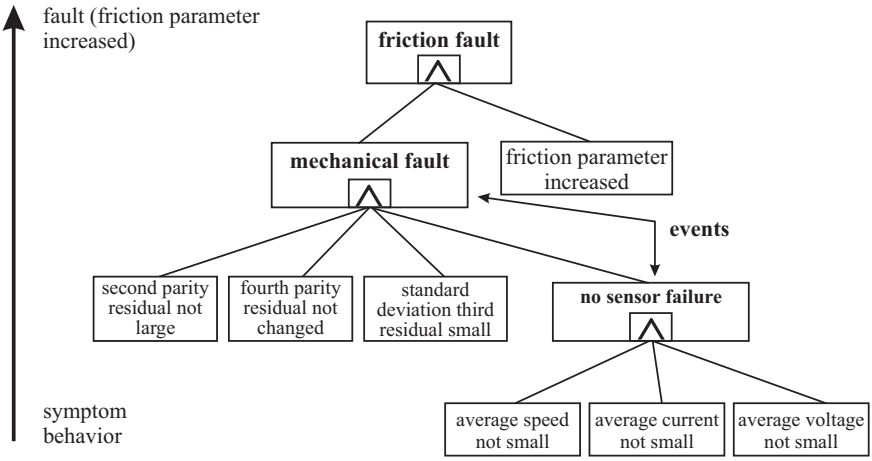


Fig. 3.7. Fault tree for one particular fault extracted from the diagnostic tree in Figure 3.5

Similar fault trees can be constructed for the other faults. This requires one to analyze the rule tree and explicitly draw the trees. The resulting set of trees is a relatively redundant representation of the fault-symptom relation because the same events are used in multiple trees. They are nevertheless very intuitive and serve to understand and visualize the functionality of the diagnostic system.

e) Computational demands

The most time-critical computation of the presented supervision concept is the computation of the continuous-time residuals. They require the evaluation of state-variable filters that are difficult to implement in fixed-point arithmetic. If the computational resources are limited, also a discrete-time form of the residuals is possible. This has, for instance, been implemented by [3.17].

The diagnosis, on the other hand, only needs to be evaluated if the fault-detection thresholds are violated. It is not time critical and can, for instance, be computed as a background job in the motor controller. Similarly, floating-point computations such as for the computation of the exponential function in the SELECT neuron can always be implemented on a lower-precision fixed-point controller, for instance, by using lookup tables. If the computational time is not critical, one can also implement floating-point arithmetic on fixed-point controllers. Since the time needed for the diagnosis is small compared with the time that typically is needed for personnel to reach a faulty device, it is obvious that the computational demand should not really be an issue. Safety-critical measures can be taken as soon as the thresholds are violated even before the diagnosis is started.

3.1.6 Conclusions

The detailed theoretical and experimental investigations with the permanently excited DC motor in idle running or with load have demonstrated that it is possible to detect 14 different faults by measurement of only three signals and combination of the parity equation and parameter estimation approach. Additive faults, like offsets of sensors, are easily detectable by parity equations in normal operation without extra input excitation signals. Multiplicative faults, like parameter deviations of the motor are better detected by parameter estimation, but require appropriate input excitation signals, at least for short times. The described methods can be transformed to other types of DC motors, depending on their construction, and also to single-phase AC motors. Further, by applying the self-learning neuro-fuzzy system SELECT all faults could be diagnosed with a 98% correct classification rate. A selection of faults, especially in the mechanical parts can also be detected by applying only signal models for current structure-borne vibrations, [3.2].

3.2 Alternating-current motor (AC)

Alternating-current motors in the form of induction or asynchronous motors consist usually of three windings placed in stator slots that are interconnected with the individual phases of a three-phase voltage supply system either in delta- or Y-connection, see Figure 3.8a). A rotating magnetic field is generated where angular velocity depends on the power supply frequency f and on the number of pole pairs p within the stator. Depending on different rotor constructions, induction motors and synchronous motors can be distinguished. In the following, induction motors with a squirrel-cage

rotor are considered. This type has a simple construction, is very robust, and is also cheap and needs less maintenance. The speed can be controlled by a field-oriented approach with variable frequency and amplitude generated by a voltage-source DC-link converter. Statistics on failure rates of AC motor show that about 50% are due to bearings, short cuts in stator windings count with 16% and broken rotor bars about 5%, [3.21], see also [3.2]. A model-based procedure for fault diagnosis of AC drives will be treated in the following, developed by [3.22].

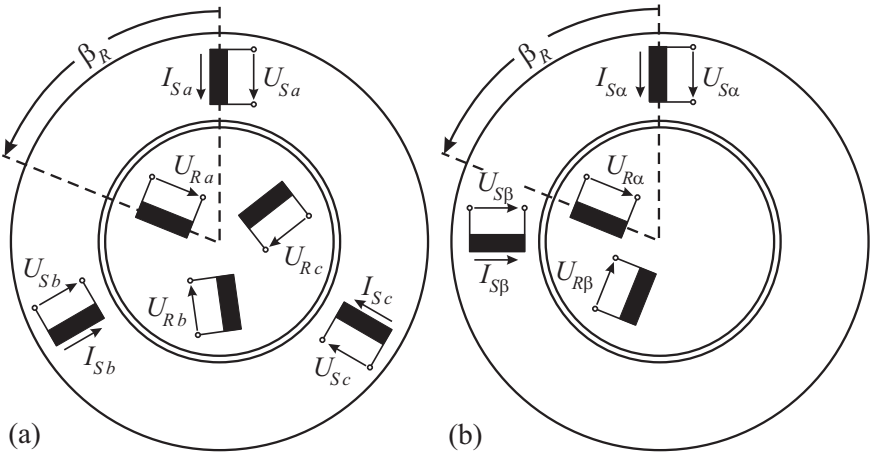


Fig. 3.8. Schematic configuration of the stator and rotor for AC motors: a) three-phase representation with one pole pair ($p = 1$) per phase; b) two-phase equivalent circuit

3.2.1 Structure and models of induction motors (asynchronous motors)

a) Electrical subsystem

Detailed models of induction motors are derived e.g., in [3.8], [3.13], [3.14]. For each rotor and stator winding the voltage and current equations are established, resulting in six coupled differential equations for a three-phase induction motor. However, by transforming the three-phase system (U_{Sa}, U_{Sb}, U_{Sc}) into a two-phase system ($U_{S\alpha}, U_{S\beta}$) via the *Clarke–Park transforms* a considerable simplification can be reached, Figure 3.8b). If the rotor flux is taken as the reference coordinate system the two-phase system is represented by (U_{Sq}, U_{Sd}), [3.8], [3.22]. Then two equations result for the rotor flux Ψ_{Rd} and electrical motor torque M_{el}:

$$T_R \frac{d\Psi_{Rd}}{dt}(t) + \Psi_{Rd}(t) = M I_{Sd}(t) \quad \text{with} \quad T_R = \frac{L_R}{R_R} \tag{3.2.1}$$

$$M_{el}(t) = \frac{3}{2} p \frac{M}{L_R} \Psi_{Rd}(t) I_{Sq}(t) \tag{3.2.2}$$

L_R	rotor self-inductance
R_R	rotor resistance
R_S	stator resistance
L_S	stator self-inductance
M	mutual inductance between stator and rotor
$I = I_{Sd} + i\,I_{Sq}$	stator current vector
p	number of pole pairs
Ψ_{Rd}	rotor flux

The flux linkage Ψ_{Rd} depends on I_{Sd} and the torque M_{el} on I_{Sq}, i.e. from each component of the stator current vector. This is the basis for the field-oriented control shown in Figure 3.9. It consists of two cascaded control loops for the flux and the speed, with the two current components as sub (minor) control variables.

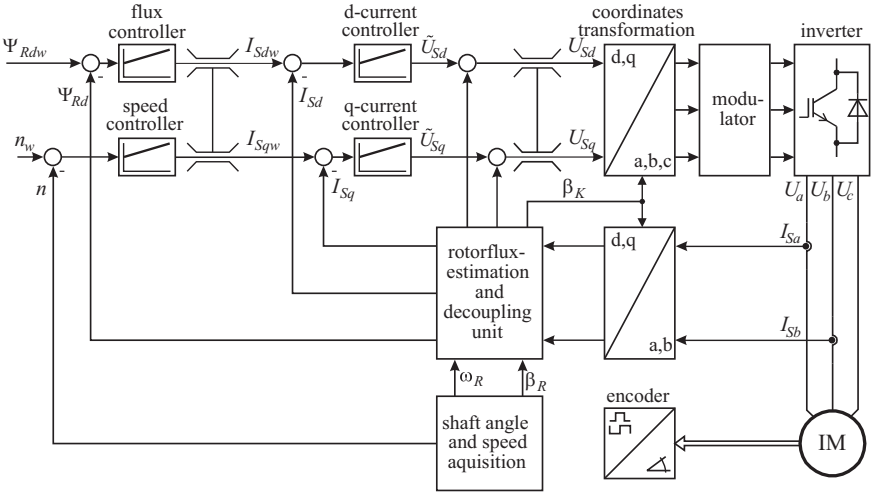


Fig. 3.9. Field-oriented control structure for induction motors

The dynamic behaviors of the electrical induction motor subsystems are

$$U_{Sd} = \left(R_S + R_R \frac{M^2}{L_R^2}\right) I_{Sd} + \sigma L_S \frac{dI_{Sd}}{dt} - \sigma L_S \omega_K I_{Sq} - \frac{R_R M}{L_R^2} \Psi_{Rd} \quad (3.2.3)$$

$$U_{Sq} = \left(R_S + R_R \frac{M^2}{L_R^2}\right) I_{Sq} + \sigma L_S \frac{dI_{Sq}}{dt} + \sigma L_S \omega_K I_{Sd} + \frac{M}{L_R} \omega_R \Psi_{Rd} \quad (3.2.4)$$

$$\sigma = 1 - \frac{M^2}{L_S L_R} \quad (3.2.5)$$

Herewith, the electrical rotor speed is $\omega_R = p\omega_m$, where ω_m is the mechanical rotor speed and the speed of the flux is ω_K with regard to the stator reference coordinate system.

b) Mechanical subsystem

The dynamic behavior of the mechanical part is obtained by establishing the angular momentum balance:

$$J\frac{d\omega_m(t)}{dt} = M_{el}(t) - M_f(t) - M_L(t) \tag{3.2.6}$$

J ratio of inertia of motor and load
M_f friction torque
M_L load torque
ω_m mechanical rotor speed

The friction torque usually consists of a Coulomb term and a viscous term:

$$M_f = M_{f0}\,\mathrm{sign}\,\omega_m(t) + M_{f1}\omega_m \tag{3.2.7}$$

The load torque depends on the connecting power consuming machine, like a pump or machine tool, and can frequently be approximated by a polynomial:

$$M_L = M_{L0} + M_{L1}\omega_m + M_{L2}\omega_m^2 \tag{3.2.8}$$

c) Thermal subsystem

Within the stator and rotor several power losses P_{LS} and P_{LR} arise which lead to a heating of the induction motor parts. The main heat sources are ohmic losses and iron losses, which can further be split up into hysteresis and eddy current losses. With the stator and rotor heat capacity

$$C_S = m_S c_{Sp}$$
$$C_R = m_R c_{Rp}$$

where m is the respective mass and c_P the specific heat value, two first-order differential equations result for the stator temperature $\vartheta_S(t)$ and rotor temperature $\vartheta_R(t)$. With further simplifications about the heat transfer through the air gap and air cooling a second-order model results for the stator temperature:

$$\Delta\vartheta_S(s) = \frac{(b_{S1}s + b_{S0})\,P_{LS}(s) + b_{R0}P_{RS}(s)}{a_2 s^2 + a_1 s + a_0} \tag{3.2.9}$$

See [3.22], [3.24].

3.2.2 Signal-based fault detection of the power electronics

Power-grid-fed converters for the supply of variable-speed AC motors frequently consist of a line-side AC–DC converter (rectifier), which rectifies the alternating grid voltage, and a motor-side three-phase DC–AC converter (inverter) that generates the three-phase system with variable frequency and amplitude, see Figure 3.10.

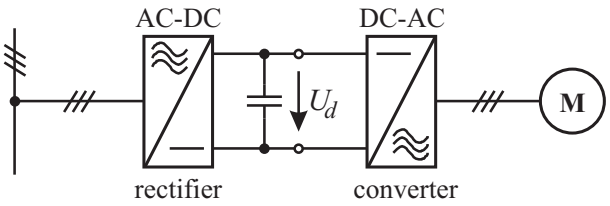


Fig. 3.10. Voltage source DC-link converter scheme for feeding AC motors

In the following it is shown how faults in these power electronics can be detected with *signal-based methods*. Measured variables are the intermediate voltage U_d, the phase currents I_{S1}, I_{S2}, I_{S3}, which are identical the I_{Sa}, I_{Sb}, I_{Sc} of Figure 3.8. For the phase voltages U_{S1}, U_{S2}, U_{S3} only the setpoints of the PWM converter are available.

a) AC–DC converter (rectifier)

Up to now, in series-produced inverters only a restricted number of diagnosis functions is implemented. The supervision functionality comprises mainly the monitoring of motor currents and the DC-link voltage as well as plausibility checks of motor parameters.

The most important faults in rectifiers are the line disconnection of one phase and faulty diodes. Among highly resistive diodes, which cannot conduct the current anymore, defective diodes also occur which have lost blocking capability in the inverse direction. This inevitably effects a phase short circuit which triggers the line fuse. However, if the fuse connected to the faulty rectifier-bridge blows, the fault is equivalent to a disconnected (open) phase. The case of highly resistive diodes is a mixture between normality and phase disconnection, as the fault has an impact only within the half period, while the faulty diode should have carried the current. The circuit diagram of the rectifier under consideration is depicted in Figure 3.11.

The faults discussed above affect high diode currents, as e.g. in the case of a disconnected phase the power has to be supplied by the remaining two phases. Although the faults do not directly lead to failure, the overloading of the remaining diodes makes an early failure most probable, [3.12], [3.23].

Significant impacts of the faults are to be observed in the run of the DC-link voltage U_d. As depicted in Figure 3.12 the signal shows significantly higher ripples in the presence of faults.

An easy and robust way for rectifier fault detection is the evaluation of the signal´s variance

$$r_{Ud} = \text{var}\ \{U_d\}\,. \tag{3.2.10}$$

As expected, the variance increases in the case of faults from the corresponding values within the healthy state. As the variance depends crucially on the load current I_{Load} it is computed for normal state in dependence on the load current, see Figure 3.13.

In order to achieve high reliability, the measured variance curve of faulty diodes is selected as the threshold. The actually measured curve exceeds the theoretical one due to noise effects and peaks arisen by the interconnected DC–AC converter. By this way faulty diodes can be detected by limit checking of the variance of the intermediate voltage.

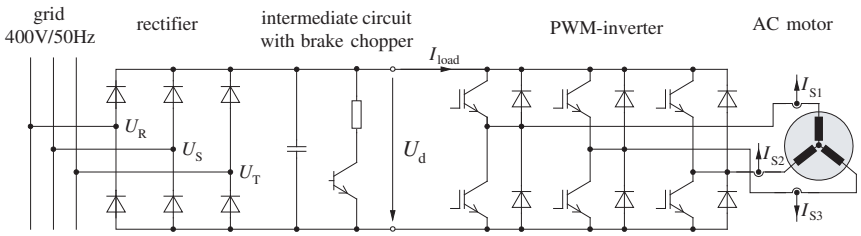


Fig. 3.11. Circuit diagram of a voltage source DC-link converter

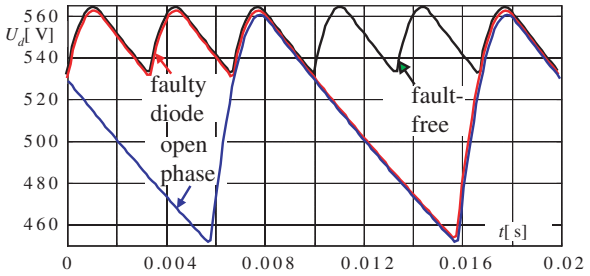


Fig. 3.12. DC-link voltage U_d in the presence of different faults (sampling frequency 2 kHz)

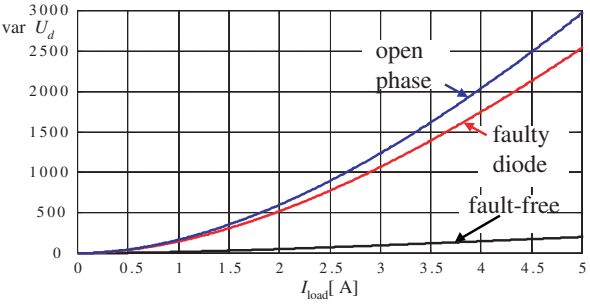


Fig. 3.13. Variance of DC-link voltage U_d for different faults

b) DC–AC converter (inverter)

Figure 3.14 shows a scheme of a considered PWM inverter. Inverter faults and also stator winding faults generate characteristic harmonics of the stator current vector. The stator current

$$\mathbf{I}_S(t) = \frac{2}{3}\left(I_{S1}(t) + I_{S2}(t)e^{-i\frac{2\pi}{3}} + I_{S3}(t)e^{-i\frac{4\pi}{3}}\right) \tag{3.2.11}$$

is transformed into an orthonormal α- and β-component coordinate system

$$\mathbf{I}_S(t) = I_{S\alpha}(t) + i\,I_{S\beta}(t) = I_{S0}(t)e^{i\varphi(t)} \tag{3.2.12}$$

In the fault-free case the trajectory of the current vector forms a circle, which deforms to an ellipse in the case of a stator winding fault, see Figure 3.15a) and to other trajectories for inverter faults and current sensor faults, Figure 3.15b)-d). In the case of these faults the spectrum of the current vector contains a positive and negative frequency, [3.22], [3.23].

$$\mathbf{I}_S(t) = \bar{I}_{S1}e^{i(\omega_S t+\varphi_1)} + \bar{I}_{S-1}e^{i(-\omega_S t+\varphi_{-1})} \tag{3.2.13}$$

where ω_S is the stator angular frequency.

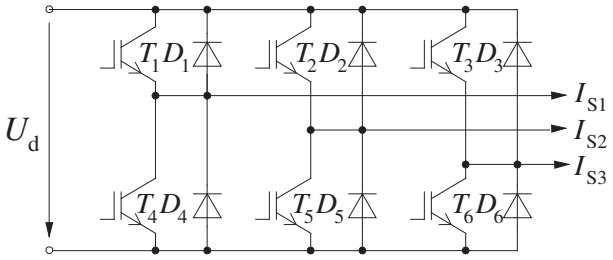


Fig. 3.14. PWM inverter for DC–AC converter

The vector trajectory $\mathbf{I}_{S-1}$ is also circular, but with smaller radius and opposite direction in the case of an ellipse. By monitoring I_{S1} and I_{S-1} obtained through a Fourier series analysis it is possible to detect the mentioned faults, [3.23]. This can be applied to AC motors with constant grid frequency f_S. However, in the case of a field-oriented control with variable frequency f_S the current vector is influenced by control dynamics. Because also the stator voltage $\mathbf{U}_S(t)$ as output of the current controllers, see Figure 3.9, shows corresponding frequencies, at least for high speeds of the motor the Fourier analysis can be performed with $\mathbf{U}_S(t)$. Therefore

$$r_{U-1}(t) = |\bar{U}_{S-1}(t)| \tag{3.2.14}$$

is taken as the fault feature for a stationary speed. However, a speed-dependent threshold is required, which has to be determined experimentally. An additional feature is the DC value

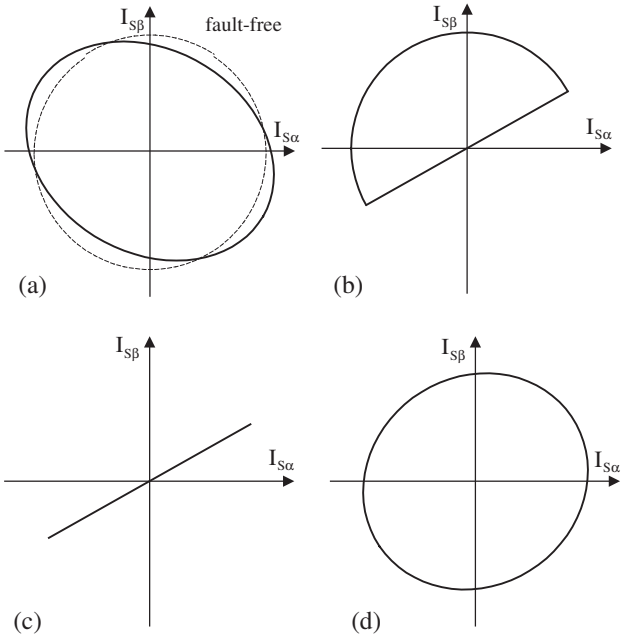


Fig. 3.15. Stator current vector trajectories: a) stator winding fault, b) inverter IGBT (Insulated Gate Bipolar Transistor) fault, c) phase 2 disconnection, d) current sensor fault

$$r_{U0Hz} = |\bar{U}_{SOHz}| \tag{3.2.15}$$

for detecting offsets in the voltage vector $\mathbf{U}_S$.

As in the case of a disconnected phase or a defective IGBT valve the three currents become different, the effective values of the three currents $I_{Si}(t)$ are calculated by taking the squares $I_{Si}^2(t)$ with subsequent low-pass filtering. Then mutual residuals are formed:

$$\begin{aligned} r_{12}(t) &= I_{S1}^2(t) - I_{S2}^2(t) \\ r_{23}(t) &= I_{S2}^2(t) - I_{S3}^2(t) \\ r_{31}(t) &= I_{S3}^2(t) - I_{S1}^2(t) \end{aligned} \tag{3.2.16}$$

In order to include all measured currents the current sum is used as a further residual:

$$r_{S0} = |I_{S0}| = |I_{S1}(t) + I_{S2}(t) + I_{S3}(t)| \tag{3.2.17}$$

which is usually zero in a fault-free situation.

Finally, Table 3.4 shows the fault-symptom relations for different faults. All considered faults can be isolated. In general, these results indicate a strong isolability. Only within the groups of open phases, defective valve and ground cuts is there a weak isolability.

Table 3.4. Fault-symptom table for the PWM converter and stator windings

Faults	Symptoms					
	$\|r_{12}\|$	$\|r_{23}\|$	$\|r_{31}\|$	$\|r_{U-1}\|$	$\|r_{UOHz}\|$	$\|I_{S0}\|$
open phase 1	++	0	++	++	0	0
open phase 2	++	++	0	++	0	0
open phase 3	0	++	++	++	0	0
defective valve 1	++	+	++	+	++	0
defective valve 2	++	++	+	+	++	0
defective valve 3	+	++	++	+	++	0
stator winding shortcut	≈ 0	≈ 0	≈ 0	+	0	0
offset fault sensor 1 or 2	≈ 0	≈ 0	≈ 0	0	+	+/++
gain fault sensor 1 or 2	≈ 0	≈ 0	≈ 0	+	0	+/++
faulty current sensor	0	0	0	0	0	+/++
ground cut phase 1	++	+	++	++	0	+/++
ground cut phase 2	++	++	+	++	0	+/++
ground cut phase 3	+	++	++	++	0	+/++

3.2.3 Model-based fault detection of the AC motor

It is assumed that the following measurements and calculated variables are available:

- U_{Sq}, U_{Sd} voltages of the q- and d-systems
- I_{Sq}, I_{Sd} currents of the q- and d-systems
- $\omega_K = \omega_S$ supply angular frequency of the PWM inverter (flux speed)
- ω_R rotor angular frequency.

In order to apply parity equations for the fault detection of the AC motor, nonlinear dynamic models are required, which are obtained by nonlinear process identification methods, [3.24].

The AC motor is the type VEM K21R90S (Normmotor) with four poles and rated values 400 V, 2.62 A, 1.1 kW, 1420 rpm (50 Hz), see [3.22].

a) Electrical part

As a basis for obtaining dynamic models (3.2.3) and (3.2.4) are used for the d- and q-subsystems. It has to be taken into account that phase voltages are not exactly known. For the practical experiments the d-current control stays closed, whereas the q-current control is opened in order to introduce an excitation signal U_{Sq}, see Figure 3.9. Therefore the rotor flux reference value Ψ_{Rdref} stays constant. Discretizing (3.2.4) with the discrete time $k = t/T_0$, where T_0 is sampling time, leads to

$$I_{Sq}(k) = \Theta_1 U_{Sq}(k) + \Theta_2 \omega_K(k) I_{Sd}(k) + \Theta_3 \omega_R(k) \Psi_{Rd} + \Theta_4 I_{Sq}(k-1) \quad (3.2.18)$$

where Θ_i are parameters, which depend on physical parameters. The product $\omega_K(k)$ $I_{Sd}(k)$ can be neglected and $\Psi_{Rd} = \text{const}$. The parameters Θ_i further depend on the operating point through ω_K and I_{Sq}. Therefore a local linear model is defined:

$$I_{Sq}(k) = w_1(\mathbf{z})U_{Sq}(k) + w_2(\mathbf{z})\omega_R(k) + w_3(\mathbf{z})I_{Sq}(k-1) \tag{3.2.19}$$

The operating-point dependence is expressed by the weighting vector

$$\mathbf{z}^T = \left[\omega_K(k) I_{Sq}(k-1)\right] \tag{3.2.20}$$

Hence, this is a semi-physical model because the structure stems from physical-based modeling. The weighting parameters $w_i(\mathbf{z})$ are estimated with the LOLIMOT identification method, [3.15], see also [3.9]. This can be considered as a special neural network, where direct least-squares parameter estimation is applied, leading to a neuro-fuzzy model.

However, the parameters depend also on the temperature of the AC motor. Therefore the stator temperature ϑ_S is measured and two correction factors $k_1(\vartheta_S)$ and $k_2(\vartheta_S)$ are introduced in (3.2.19):

$$I_{Sq}(k) = w_1(\mathbf{z})k_1(\vartheta_S)U_{Sq}(k) + w_2(\mathbf{z})k_2(\vartheta_S)\omega_R(k) + w_3(\mathbf{z})I_{Sq}(k-1) \tag{3.2.21}$$

These correction factors are estimated and $k_1(\vartheta_S)$ is approximated by a second-order polynomial and $k_2(\vartheta_S)$ with a linear dependence.

The dynamic behavior of the AC motor was identified by excitation of U_{Sq} with an APRBS, an amplitude-modulated PRBS, and sampling time $T_0 = 1.5\,\text{ms}$ (667 Hz). The obtained generalization results in Figure 3.16 show a very good agreement with six local models and two correction characteristics for stator temperatures $\vartheta_S \in [25°\text{C}, 60°\text{C}]$.

In a similar way the d-system can be identified. (3.2.3) and experimental trials lead to

$$U_{Sd}(k) = w_0(\mathbf{z}) + w_1(\mathbf{z})\omega_K(k) + w_2(\mathbf{z})I_{Sq}(k) \tag{3.2.22}$$

As I_{Sd} is constant, its derivative is zero. Therefore the d-model is static. The temperature dependence is again considered by correction factors. Figure 3.17 shows relatively good results with a model having six local linear models and 18 correction characteristics.

These nonlinear precise models can now be used to apply parity equations for fault detection. The following output residuals and their variances are formed, compare Figure 3.18:

$$r_q = I_{Sq} - \hat{I}_{Sq} \tag{3.2.23}$$

$$r_d = U_{Sd} - \hat{U}_{Sd} \tag{3.2.24}$$

furthermore r_{12}, r_{23}, r_{31} and $|I_{S0}|$ form the phase currents, (3.2.16) and (3.2.17). As the models are more precise in the case of stationary behavior, an adaptive threshold is used for dynamic states in dependence on the current I_{Sq}, which is proportional to the torque-generating dynamics. Table 3.5 presents the fault-symptom relation for different faults. The AC motor faults stator winding defect, broken rotor bar, and rotor eccentricity are strongly isolated and can therefore be diagnosed. However, broken rotor bar and broken end ring are only weakly isolated and cannot be clearly separated. For the other faults the same results are obtained as in Table 3.4 for the PWM converter.

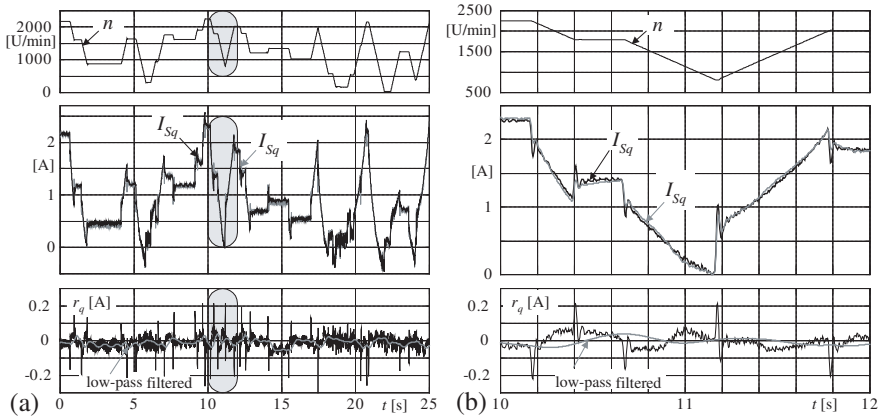


Fig. 3.16. Generalization data for LOLIMOT identification of the q-system: a) Input: U_{Sq} (APRBS). Output: I_{Sq}, n and output error $r_q = I_{Sq} - \hat{I}_{Sq}$ for $\delta_S = 25°C$; b) zoomed signals

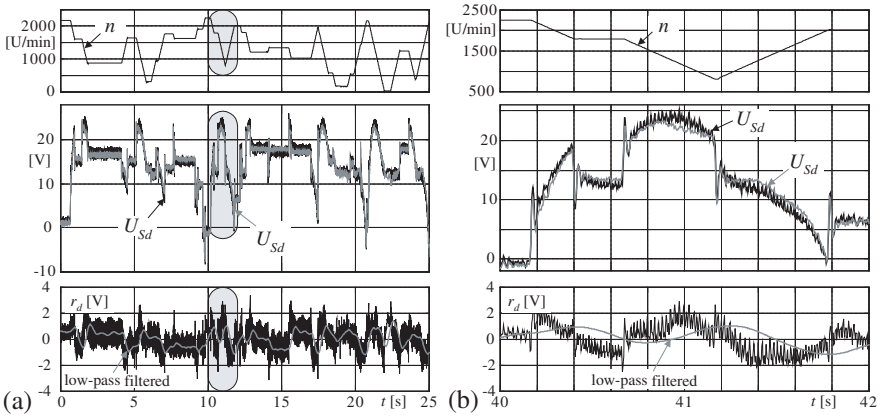


Fig. 3.17. Generalization data for LOLIMOT identification of the d-system: a) Input: U_{Sq} (APRBS). Output: U_{Sd}, n and output error $r_d = U_{Sd} - \hat{U}_{Sd}$ for $\delta_S = 25°C$; b) zoomed signals

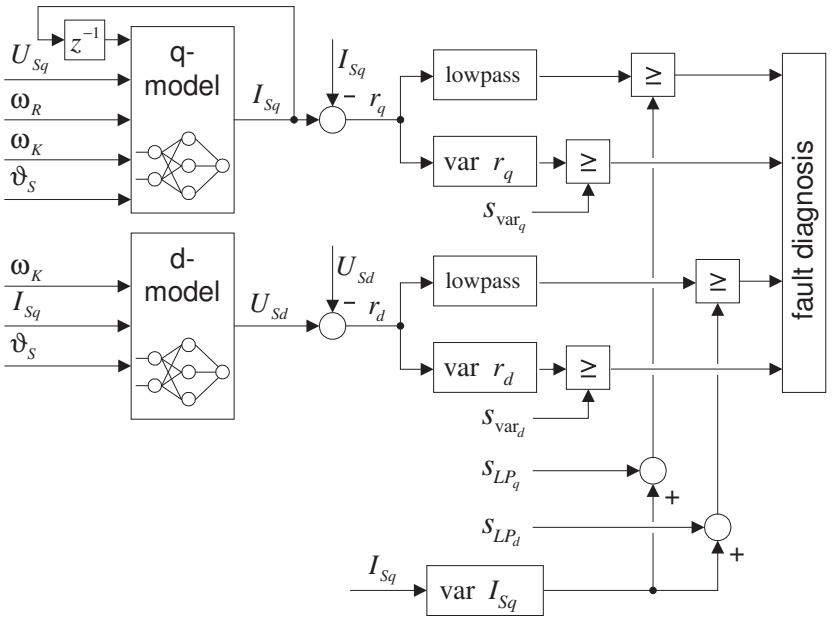


Fig. 3.18. Residual generation for the AC motor with nonlinear parity equations, S: thresholds

Table 3.5. Fault-symptom table for the AC motor

Faults	Symptoms							
	$\lvert r_{qLP}\rvert$	var $\{r_q\}$	$\lvert r_{dLP}\rvert$	var $\{r_d\}$	$\lvert r_{12}\rvert$	$\lvert r_{23}\rvert$	$\lvert r_{31}\rvert$	$\lvert I_{S0}\rvert$
open phase 1	−−	++	++	++	++	0	++	0
open phase 2	−−	++	++	++	++	++	0	0
open phase 3	−−	++	++	++	0	++	++	0
defective valve 1	−	+	++	++	++	+	++	0
defective valve 2	−	+	++	++	++	++	0	0
defective valve 3	−	+	++	++	+	++	++	0
stator winding shortcut	−−	0	+	+	≈ 0	≈ 0	≈ 0	0
broken rotor bar	++	+	+	++	0	0	0	0
broken end ring	+	+	+	+	0	0	0	0
rotor eccentricity	+	0	0	0	0	0	0	0
gain fault sensor 1 or 2	−/+	0	0	0	0	0	0	+/++
offset sensor fault 1 or 2	+	0	0	0	0	0	0	+/++
fault current sensor	0	0	0	0	0	0	0	+/++
ground cut phase 1	−−	+	+	+	++	+	++	+/++
ground cut phase 2	−−	+	+	+	++	++	+	+/++
ground cut phase 3	−−	+	+	+	+	++	++	+/++

b) Mechanical subsystem

The dynamic behavior of the rotor speed $\omega_R(t)$ follows from (3.2.6). Faults in the mechanical part express themselves especially in the friction parameters M_{f0} and M_{f1} and eventually in the ratio of inertia J. However, these parameters also depend on the connected load like general drive-trains with or without gears and power-consuming machines like machine tools or pumps. Therefore this dynamics equation depends on the load and corresponding available measurements, see e.g. [3.9]. The electrical torque M_{el} can be determined with (3.2.1) and (3.2.2) using the current I_{Sd} and I_{Sq} of the d- and q-systems, which are known within a field-oriented controller, and the mechanical rotor speed ω_m. Figure 3.19 summarizes the resulting signal flow and computations. Some electrical parameters or groups of parameters can then be estimated. As the mechanical subsystem is slower than the electrical subsystem, a larger sampling time can be chosen, e.g. $T_0 = 10\,\text{ms}$ (100 Hz).

An application with a circular pump is reported in Section 6.1.

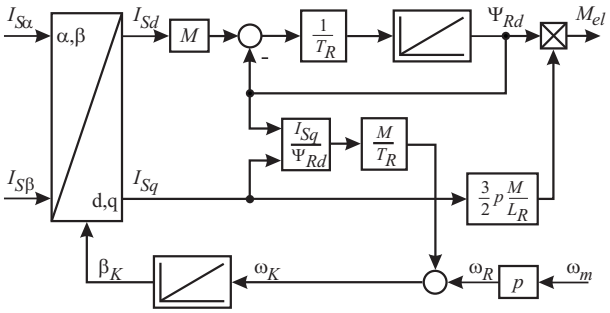


Fig. 3.19. Calculation of the electrical torque M_{el} based on measurements of I_{Sd}, I_{Sq} and ω_m

c) Thermal subsystem

The thermal state of the AC motor is indicated by the temperature of the rotor and the stator. Overheating arises because of defective cooling, high friction and overload. The generated stator heat power due to ohmic losses is

$$P_{LS} = \frac{3}{2} R_S \left(I_{Sd}^2 + I_{Sq}^2 \right) \tag{3.2.25}$$

and the rotor losses. If the rotor power losses are neglected, only one part of the transfer function of (3.2.9) has to be considered. The corresponding z-transfer function

$$G(z) = \frac{\Delta\vartheta_S(z)}{P_{LS}(z)} = \frac{\beta_1 z^{-1} + \beta_2 z^{-2}}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2}} \tag{3.2.26}$$

possesses two poles which belong to a large time constant $T_1 \approx 17\,\text{min}$ and a small one $T_2 \approx 2\,\text{min}$, which were determined by parameter estimation (sampling time $T_0 = 30\,\text{s}$). Then an output residual as the difference of the measured and estimated stator temperature

$$r_\vartheta = \vartheta_S - \hat{\vartheta}_S \tag{3.2.27}$$

can be formed. In the case of dynamic operations parameter estimation can be applied. Especially the time constant $\hat{T}_1$ then indicates faults of the thermal system.

Figure 3.20 shows the increase of the stator temperature for a defective fan wheel and Table 3.6 the estimates of the time constants. For cooling faults or overload the stator temperature increases considerably, indicated by a larger gain of (3.2.26). Also the large time constant T_1 shows a strong increase. The small constant T_2 remains approximately constant.

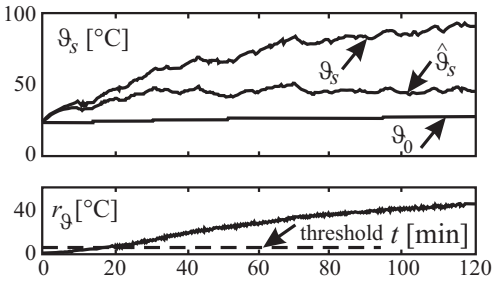


Fig. 3.20. Stator temperature ϑ_S and temperature residual r_ϑ in the case of a defective fan wheel. $\hat{\vartheta}_s(P_{LS}, t)$ is the model output for normal behavior

Table 3.6. Parameter estimates of the thermal subsystem for different cooling defects

time constants	fault-free	defective fan wheel	covered cooling slots	covered motor
$\hat{T}_1$ [min]	16.6	69.3	31.2	36.1
$\hat{T}_2$ [min]	1.99	1.86	1.82	1.72

d) AC motor at standstill

Some faults in induction motors do not immediately effect a complete failure of the drive. Thus, e.g. in the case of broken rotor bars further operation possible in principle. In consideration of the higher currents in the adjacent bars and the resulting mechanical stresses due to thermal overload and unbalance, further rotor or end-ring breakage may occur [3.21].

Hence, the monitoring of these faults does not have to be performed permanently. Basically, it is sufficient to supervise the induction motor after certain periods of time.

In order to select an adequate operating point for the test of the electrical subsystem some basic considerations are necessary. Usually the motor shaft is linked to a specific load (e.g. centrifugal pump) and any operating point with motor speeds different from zero affects the load. If the monitoring approach must not influence the load facility and the disconnection of the load for this purpose is undesirable, the only possible solution is to perform the supervision cycle at *standstill*. The background is that the feeding voltage of the induction machine can be generated by the inverter in a manner that the motor does not produce any torque. Furthermore, the motor speed is not required and does not have to be measured.

For the test signal different selections can be made to cover a certain frequency range. On the one hand *sinusoidal input signals* can be chosen. With the help of the measured output currents the impedance is determined for several frequency points by employing frequency-response methods. Subsequently, the physical parameters (rotor/stator resistances/inductances, mutual inductance) are estimated by means of least-squares parameter estimation methods. On the other hand, the process can also be excited by PRBS (pseudo-random binary signals). Here, the physical parameters are determined by estimating a dynamic continuous single input/output model with non-recursive least-squares parameter-estimation methods, see Figure 3.21, [3.1].

The parameters are estimated for different excitation axes and in the presence of specific faults characteristic curves are obtained. Both described approaches have been implemented and tested and similar estimation results have been obtained. The results achieved with the frequency-response approach are more precise. On the other hand, the second approach is approximately ten times as fast as the frequency-response technique.

In the stationary case the AC motor can be considered as a transformer, [3.22]. The transfer function then results as

$$G_{Sd}(s) = \frac{I_{Sd}(s)}{U_{Sd}(s)} = \frac{1}{R_S} \frac{1 + T_R s}{1 + (T_R + T_S)s + T_R T_S \sigma s^2} \tag{3.2.28}$$

with $T_S = L_S/R_S$ and $T_R = L_R/R_R$. This leads to the differential equation

$$\ddot{I}_{Sd}(t) = b_0 U_{Sd}(t) + b_1 \dot{U}_{Sd}(t) - a_0 I_{Sd}(t) - a_1 \dot{I}_{Sd}(t) \tag{3.2.29}$$

with

$$b_0 = \frac{1}{\sigma T_S T_R R_S} \quad b_1 = \frac{1}{\sigma T_S R_S} \quad a_0 = \frac{1}{T_S T_R \sigma} \quad a_1 = \frac{T_S + T_R}{T_S T_R \sigma}$$

With the help of parameter estimation and state-variable filters for determining the derivatives it is then possible to estimate $\hat{R}_S$, $\hat{L}_S$, $\hat{R}_R$ and $\hat{M}$ and to detect faults like a broken bar, broken end ring and eccentricity. Figure 3.21 shows the measurement configuration and Figure 3.22 depicts the increased rotor resistance in dependence on the excitation axis for a broken bar. More details are given in [3.22].

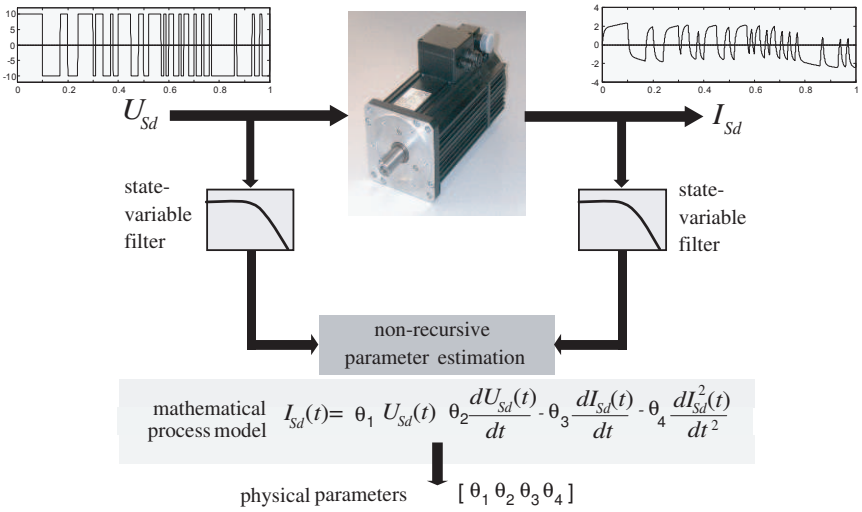


Fig. 3.21. Scheme for the parameter estimation of stator and rotor parameters at standstill

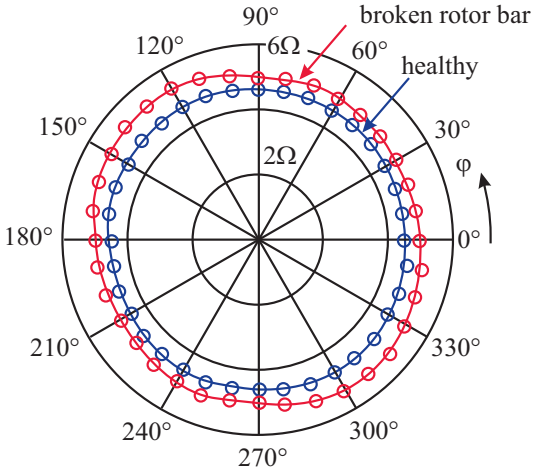


Fig. 3.22. Estimated rotor resistance R_R in dependence on the excitation axis φ

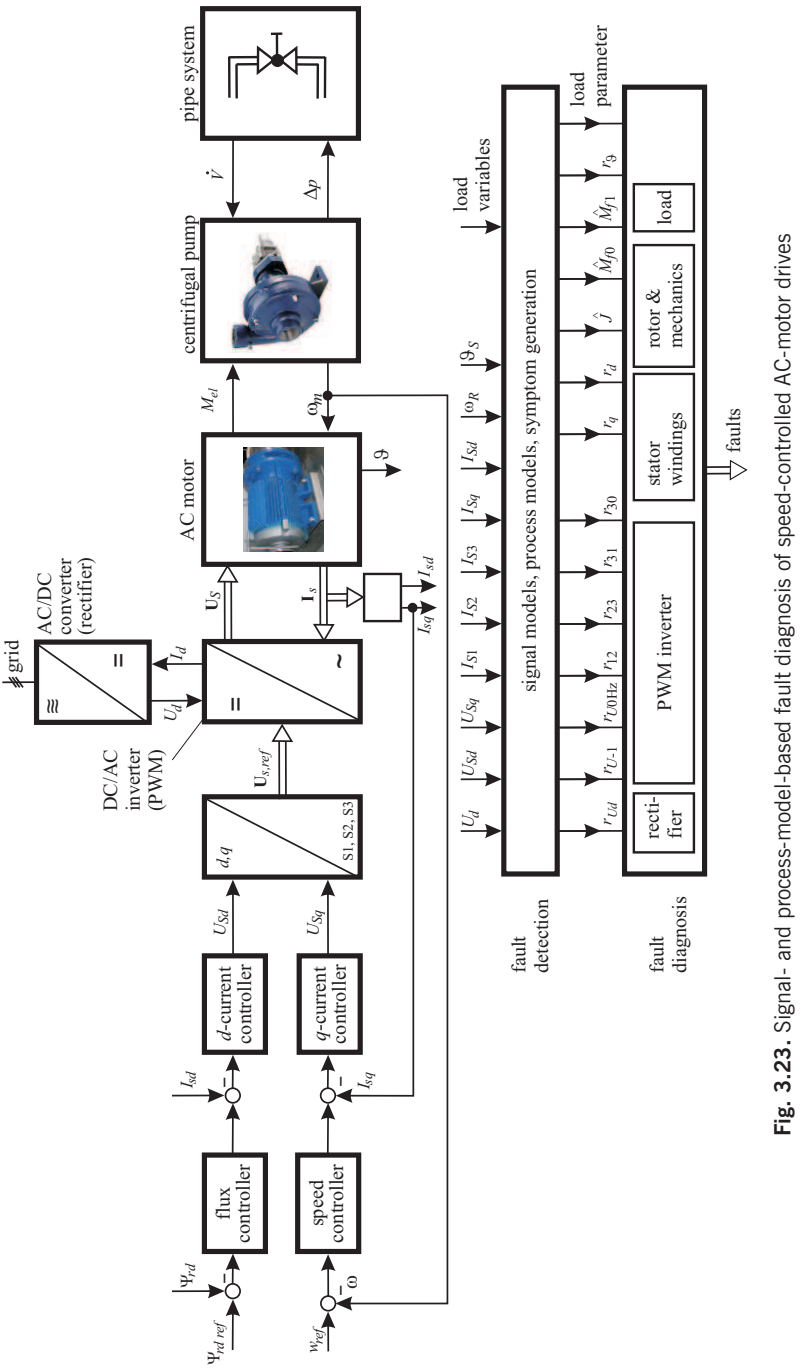


Fig. 3.23. Signal- and process-model-based fault diagnosis of speed-controlled AC-motor drives

3.2.4 Conclusions

Figure 3.23 gives an overall scheme of the signal- and process-model-based fault detection and diagnosis of a speed-controlled AC drive. The investigations have shown that with the support of physical modeling it is possible to detect several different faults in the rectifier, the PWM inverter, the AC-motor stator and rotor, and the motor mechanics. The use of nonlinear output *parity equations* is especially attractive for the electrical subsystems. However, it required relatively precise process models. These models can be obtained with a multi-model approach by applying local linear models and the LOLIMOT *nonlinear parameter estimation*. For the fault detection of the motor mechanics *linear parameter estimation* is preferable and includes also process parameters from the load machinery. The model-based fault-diagnosis approach requires only four sensors and some variables which are available within the field-oriented control, determines up to 14 symptoms and can diagnose about 10 different faults. The described fault-detection methods can be directly transferred to synchronous motors.

4

Fault diagnosis of electrical actuators

Actuators usually transform low-powered manipulated variables (e.g. analog voltages 0−−10 V, applied DC currents 0−−20 mA or 4−−20 mA, pneumatic pressures 0.2 − −1 bar, or hydraulic pressures 0 − −150 bar) into process input variables of a much higher power level. Frequently the process input variable is a flow of energy or matter, or a force or torque. The power needed for actuating is provided by an auxiliary energy supply, which feeds the power amplifier for the actuator. The auxiliary energy can be electrical, pneumatic or hydraulic. In many cases the actuators are composed of a signal transformer, an actuator drive, an actuator transformer (gear, spindle) and an actuating device or valve, compare Figure 4.1 and [4.7]. Actuators can operate in open loop or closed loop (e.g. position or flow-control). A survey of basic structures of actuators, different types, characteristics and mathematical models is given in [4.4].

Actuators play an important role in any manually or automatically controlled system. In the following sections the fault detection of electrical and fluidic actuators is considered.

4.1 Electromagnetic actuator

Electromagnets play an important role as actuators in many technical processes. Therefore their supervision and fault detection is of primary interest. One distinguishes switching magnets and proportional magnets. The *switching magnets* have the task to switch, e.g. a valve from one end position to the other end position. Therefore, only switching and reaching of the end positions are of interest. *Proportional magnets* have to hold any position between the two end positions. They act usually against a spring, compare Figure 4.2. Because of the required precise positioning they should have a linear characteristic between the input voltage U and the magnetic force F and small electrical and mechanical (frictional) hysteresis. Usually, the proportional actuator is operating in a closed loop for position control, where the controlled variable is the measured position of the armature or another output variable like pressure. To design the force characteristic $F(z)$ independently of the air

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_4, © Springer-Verlag Berlin Heidelberg 2011

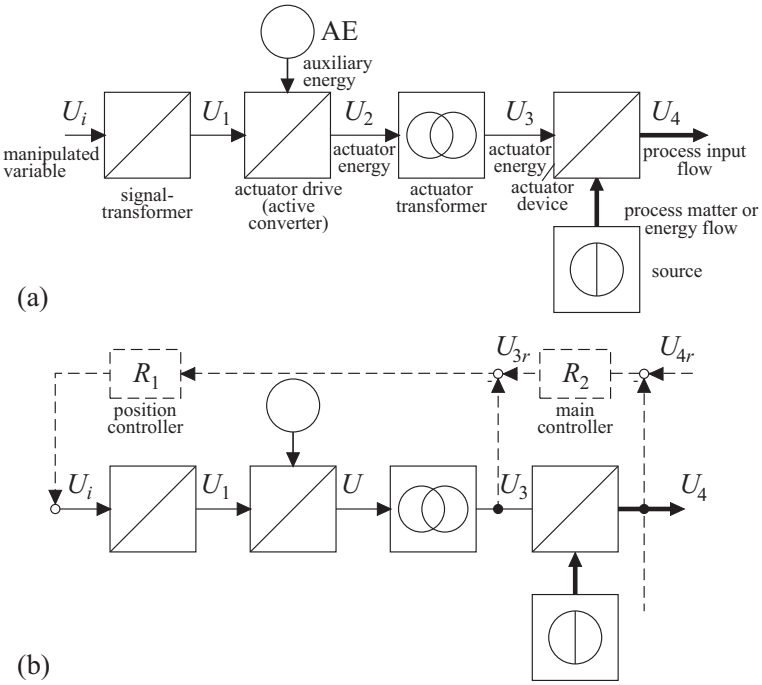


Fig. 4.1. Basic structures of actuators: a) open-loop controlled actuator; b) closed-loop controlled actuator

gap the magnetic yoke is frequently designed conically around the operating point of the armature, [4.4], [4.8], see Figure 4.2. The nonlinear force–position characteristic of a switching magnet, which has no constant characteristic $F(z)$, can also be linearized by a nonlinear compensation as shown in Figure 4.3.

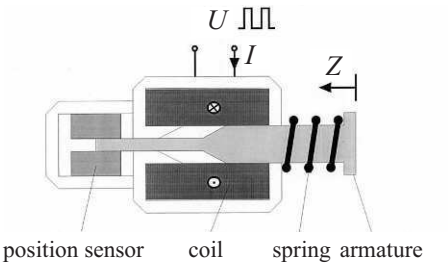


Fig. 4.2. Scheme of the investigated low-cost DC solenoid drive

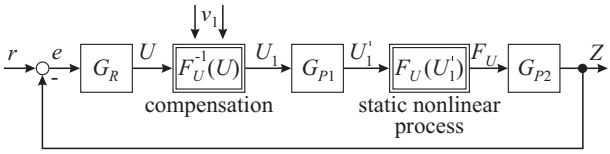


Fig. 4.3. General structure of a series correction (compensation) for nonlinear process statics

4.1.1 Position control

As an example a simple solenoid designed as in Figure 4.2 is considered with the following data, [4.16]:

armature length	125 mm
armature diameter	25 mm
coil length	60 mm
coil resistance	$R = 22.4\,\Omega$
inductivity	$L(Z = 0)$ mm = 0.87 H
	$L(Z = 25)$ mm = 1.18 H
voltage	$U = 24$ V (DC)
spring constant	$c_F = 1620$ N/m
position sensor	inductivity, accuracy 0.5%
	measurement range: 40 mm
	time constant: 2.5 ms

An appropriate function for describing the nonlinear force–current characteristic, [4.8], see Figure 4.4, is the polynominal approximation

$$F(I, Z) = I \sum_{i=0}^{2} \frac{K_i}{(Z_0 - Z)^i} \quad \text{with} \quad Z_0 = 26\,\text{mm} \tag{4.1.1}$$

The resulting statics of the linearized actuator are shown in Figure 4.5 where a typical hysteresis characteristic becomes obvious. Its gradient represents the local gain K_P of the actuator, which can after linearization be assumed constant. The position-dependent width of the hysteresis characteristic is a measure for frictional forces and magnetic hysteresis.

In addition to the compensation of the nonlinear force characteristic an adaptive friction compensation can be applied, based on the parameter estimation of the Coulomb-friction coefficient, as shown in [4.4], [4.6] and [4.16].

The linearized system including the compensation of the nonlinear force–position characteristic can now be described by two equations. For the inner current circuit

$$T_I \dot{I}(t) + I(t) = K_I U(t) \tag{4.1.2}$$

holds approximately and for the mechanical subsystem with $\Delta Z = z$

$$m\ddot{z}(t) + d\dot{z}(t) + cz(t) = K_{Mag} I(t) - F_C \operatorname{sign}(\dot{z}) + F_L(t) \tag{4.1.3}$$

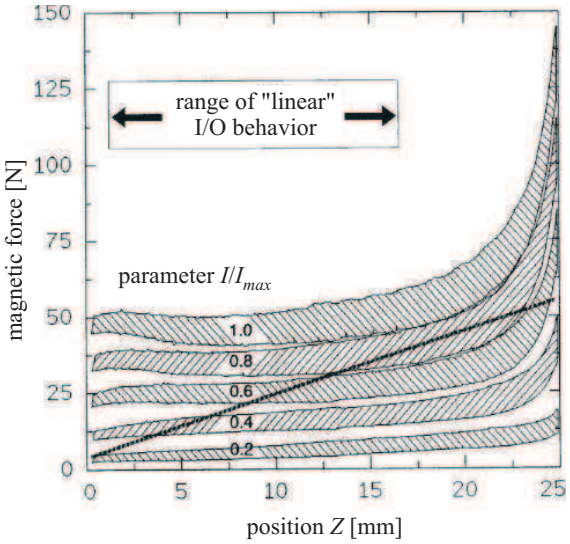


Fig. 4.4. Position-dependent nonlinear force–position characteristic of the (switching) solenoid drive. The straight line represents the linear spring characteristic

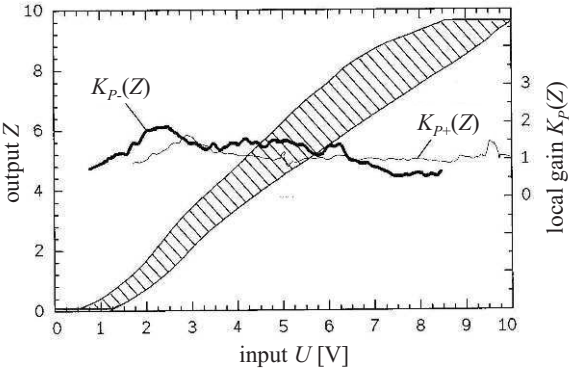


Fig. 4.5. Hysteresis characteristic and position-dependent local gain of the *linearized* (compensated) solenoid drive (1 V $\triangleq$ 2.5 mm)

The I/O behavior of the actuator can be modeled as a third-order system. The unknown parameters are obtained during a pre-identification phase, exciting the actuator with a special input signal with sampling time $T_0 = 2.5\,\text{ms}$. Taking the effect of Coulomb friction into account, the output error parameter estimation leads to the following direction-dependent transfer functions:

$$G_+(s) = \frac{z(s)}{U(s)} = \frac{382.4}{(s+116.4)\left(s^2+40.4\,s+3329.4\right)} e^{0.0025\,s} \tag{4.1.4}$$

$$G_-(s) = \frac{z(s)}{U(s)} = \frac{220.0}{(s+47.9)\left(s^2+47.9\,s+3444.5\right)} e^{0.0025\,s} \tag{4.1.5}$$

The index $+/-$ denotes the direction of the armature motion and the additional dead-time describes the effect of asynchronous PWM generation.

Figure 4.6 shows the obtained control performance using a numerical optimized position controller

$$G_q\left(q^{-1}\right) = \frac{\Delta U(k)}{r(k)} = \frac{2.231 - 4.204q^{-1} + 2.000\,q^{-2}}{\left(1-q^{-1}\right)\left(1-0.616q^{-1}\right)} \tag{4.1.6}$$

proportional-integral derivative (PID) type with first-order lag, $T_0 = 2.5\,\text{ms}$ where q^{-1} is a shift operator for one sampling time $u(k)q^{-1} = u(k-1)$. Although there is a change in the actuator´s dynamic behavior, the controller designed for the slower negative motion (worst case) is robust enough for positive motions. The dynamic features are suitable and stability is obtained even in the positioning range $17\,\text{mm} < Z < 25\,\text{mm}$, which is unstable with linear control and uncompensated actuator characteristic.

4.1.2 Fault detection with parameter estimation

The electromagnetic actuator is now considered for the linearized operation range $0 - 25\,\text{mm}$. From the equations for the current circuit (4.1.2) and the mechanical subsystem (4.1.3) a third-order differential equation follows:

$$z^{(3)}(t) + a_2^*\ddot{z}(t) + a_1^*\dot{z}(t) + a_0^* z(t) = b_0^*\Delta U(t) + c_0^*(t) \tag{4.1.7}$$

The parameters of the continuous-time representation

$$\boldsymbol{\Theta}^T(t) = [a_2^* a_1^* a_0^* b_0^* c_0^*] \tag{4.1.8}$$

depend thereby on the physical process coefficients

$$\mathbf{p}^T = [T_1, D, \omega_0, K_P, c_0^*] \tag{4.1.9}$$

with, for example, [4.4],

$$D = \frac{d}{2\sqrt{mc}}, \omega_0 = \sqrt{\frac{c}{m}} \tag{4.1.10}$$

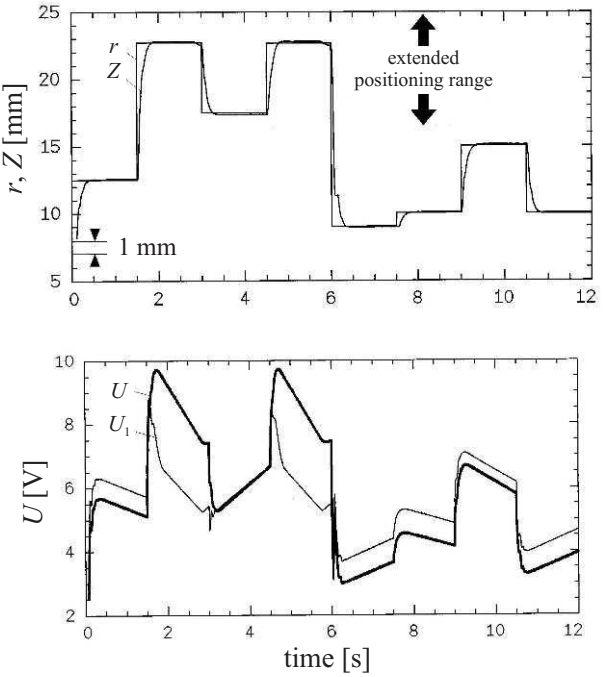


Fig. 4.6. Closed-loop position control of the solenoid with correction of the nonlinear actuator characteristic, but without friction compensation, $T_0 = 2.5$ ms. U_1; after compensation of the nonlinearity, Figure 4.3

These process coefficients can be expressed in terms of the parameter estimates $\boldsymbol{\Theta}$. Hence, after estimation of the model parameters $\boldsymbol{\Theta}$ by measuring the voltage U and the position Z all process coefficients $\mathbf{p}$ can be calculated, [4.16].

In the following some experimental results are shown for artificially generated actuator faults:

F1: too large of the spring pre-tension
F2: decrease of the spring constant (by break or aging, change from c = 1650 to 1200 N m^{-1})
F3: increase of friction (increase of surface roughness and jamming)
F4: fault in the current circuit (weak controller gain).

The parameters were estimated by output error minimization using specific excitation signals. Sampling time was $T_0 = 0.2$ ms. Figure 4.7 and Table 4.1 show the results for different faults. Based on the deviations (symptoms) all faults can be identified. This can be performed by a pattern recognition or a systematic treatment of fault-symptom trees. In all cases different patterns of coefficient changes result. This enables a unique diagnosis of the four faults based on parameter estimation.

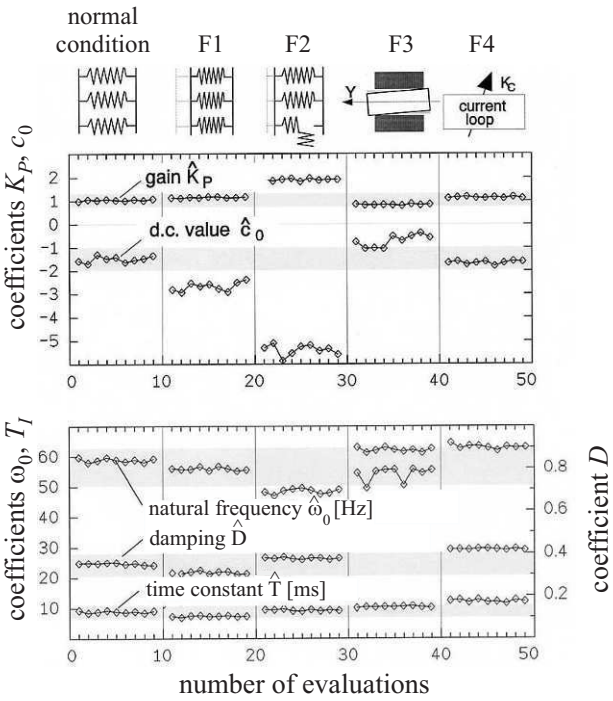


Fig. 4.7. Parameter estimates for an electromechanical drive with different faults (positive motion direction)

Table 4.1. Changes of process coefficients for an electromagnet depending on different faults

	static coefficients		dynamic coefficients		
Fault type	K_{P+}	c_{0+}	ω_{0+}	D_+	T_1
F1	0	−−	0	0	0
F2	++	−−	−	+	0
F3	−	+	0	++	0
F4	0	0	0	+	++

4.2 Electrical automotive throttle valve actuator

Since about 1990, electrically driven throttle valves became a standard component for gasoline engines. They control the air mass flow through the intake manifold to the cylinders. The electrical throttles are manipulated by the accelerator-pedal sensors via an electronic control unit and additional control inputs from idle-speed control, traction control and cruise control. In many vehicles it is the first drive-by-wire component, replacing the former mechanical linkage, [4.17]. Because the electrical throttles are safety-related components, reliability and safety are of high importance. In the following a fault detection and diagnosis method and the corresponding equip-

ment is described which was developed for the computer-based testing of electrical throttle components for quality control. The methods can also be applied onboard.

4.2.1 Structure and models of the actuator

Figure 4.8 shows a cross-sectional view of the actuator. A permanently excited DC motor with brush commutation drives the throttle through a two-stage gear in the opening or closing direction. It operates against the main helic spring. A second spring works in the closing region in the opposite direction, in order to open the throttle in the case of voltage loss into a limp-home position (a mechanical redundancy). The motor is manipulated by a pulse-width-modulated (PWM) armature voltage U_A(−12 to + 12 V). The measured variables are the armature voltage U_A, the armature current I_A and the angular throttle position φ_K(0 to 90°). This throttle position is measured by two redundant wiper-potentiometers operating in two different directions. Some technical data are given in Table 4.2. The position controller was a model-based sliding-mode controller or PID controller with time lag and sampling time $T_0 = 1.5$ ms, [4.14], [4.15].

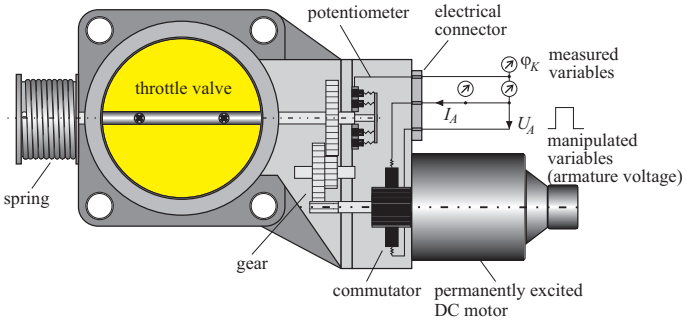


Fig. 4.8. Scheme of the electrical throttle

Theoretical modeling of the throttle valve leads to the following basic equations, compare (3.1.1), (3.1.2) and [4.4]:

Electrical part,

$$U_A(t) = R_A I_A(t) + \Psi\omega_A(t) + c_{0e} \tag{4.2.1}$$

$$M_{el}(t) = \Psi I_A(t) \tag{4.2.2}$$

Mechanical part (related to motor axle),

$$\nu J\dot{\omega}_k = M_{el}(t) - M_{mech}(t) \tag{4.2.3}$$

$$M_{mech}(t) = \frac{1}{\nu}\left(c_{S1}\varphi_k(t) + M_{S0} + M_F\right) \quad (\varphi_k > \varphi_{k0}) \tag{4.2.4}$$

$$M_F(t) = M_{F0}\,\mathrm{sign}\,\omega_k(t) + M_{F1}\omega_k(t) \tag{4.2.5}$$

The signals used are:

R_A	armature resistance
Ψ	magnetic flux linkage
ν	gear ratio ($\nu = 16.42$)
J	moment of inertia of the motor
M_{F0}	Coulomb-friction torque
M_{F1}	viscous-friction torque
c_{S1}	spring constant
M_{S0}	spring pre-tension
$\omega_k = \dot{\varphi}_k$	throttle angular speed
ω_A	motor angular speed. $\omega_A = \nu\omega_k$

Table 4.2. Technical data of the electromagnetic throttle actuator (Bosch, DV-E4). Permanently excited DC motor with brushes. 1 pole pair, 12 commutation segments, 2 ball bearings; 12 V, 1.4 A, reversion current: 7.5 A; throttle diameter: 70 mm, needle bearings

Parameter			Reference value
armature resistance	R_A	[Ω]	1.2
inductance	L_A	[μ H]	600
magnetic flux linkage	Ψ	[Nm/A]	0.029
moment of inertia of the motor	J	[kg m^2/rad]	0.0000092
spring constant	c_{S1}	[Nm/°]	0.002 – 0.0021
spring pre-tension	M_{S0}	[Nm]	0.29 ± 0.03
Coulomb-friction torque	M_{F0}	[Nm]	ca. 0.18
gear ratio	ν	[-]	43/12*55/12 = 16.42

Compare the general equations for a DC motor in Section 3.1 and Figure 4.9. The armature inductance can be neglected, because the electrical time constant $T_{el} = L_A/R_A \approx 1$ ms is much smaller than the mechanical dynamics. The constant c_{0e} takes additive faults into account.

Dependent on the input excitation either the Coulomb friction or the viscous friction turned out to be dominant. Figure 4.10 depicts the principle of the applied fault detection and diagnosis.

4.2.2 Input test cycle for quality control

In order to achieve optimal diagnostic results, the test cycle is composed of different phases, which each phase is designed to get a deep insight into a specific subsystem or technical component of the whole throttle valve actuator, see Figure 4.11.

At the beginning of the test (phases 1––4) the throttle valve actuator is controlled in an open loop by directly manipulating the armature voltage of the DC motor. In this first step, an open circuit, a short circuit or a sneak path between the electrical lines of the actuator are detected. Furthermore, offset faults in the measured signals armature voltage U_A, armature current I_A and throttle position φ_K are detected.

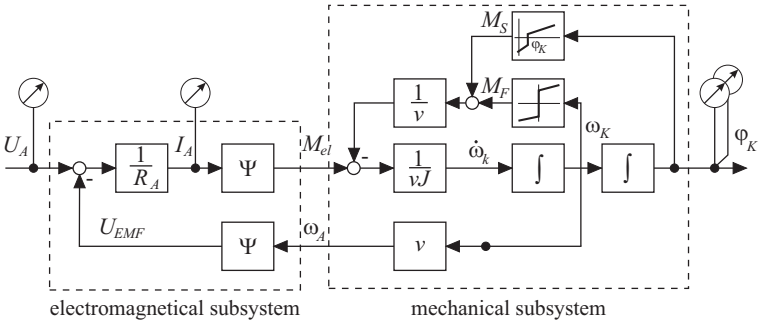


Fig. 4.9. Signal flow diagram of the electromagnetic throttle actuator

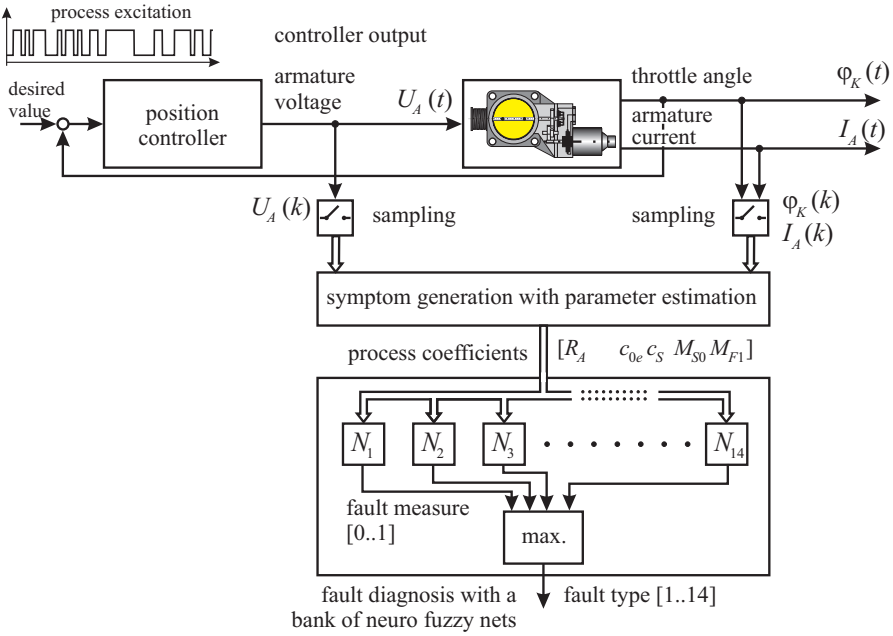


Fig. 4.10. Schematic of the fault-diagnosis system of the electromagnetic throttle actuator. N: normalization

After this first stage the diagnosis proceeds in the closed-loop mode (phases 5 and 6). Now the manipulated variable of the diagnostic algorithm is given by the throttle position setpoint. This stage consists of two phases. In the first one, the whole range of operation is covered by a triangular setpoint. Herewith the redundantly measured position signals are checked for plausibility, and an insight into the mechanics of the actuator is gained by the estimation of some parameters of the mechanical subsystem. In phase 6, the test object is excited with a high dynamic signal in order to achieve a large variation of the speed of the DC motor. Then the parameters of the electrical subsystem of the motor are determined by continuous-time parameter-estimation algorithms. Furthermore, dynamic deviations from the nominal process behavior are detected with a parity equation using the model obtained from parameter estimation.

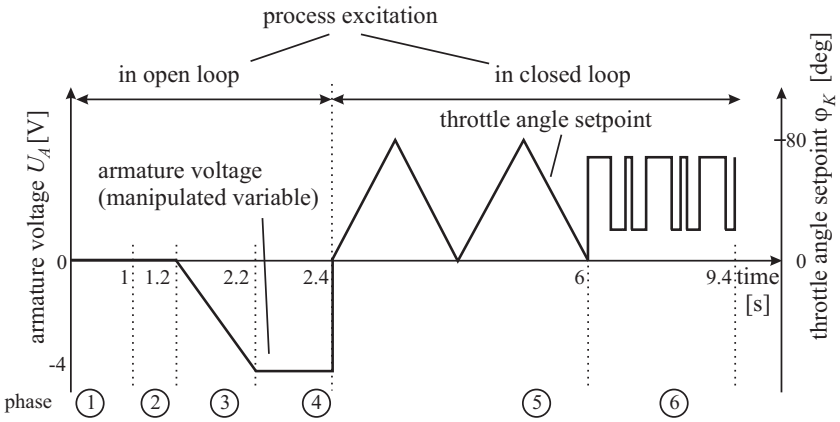


Fig. 4.11. Test cycle for the automatic fault diagnosis of the throttle valve

4.2.3 Fault detection with parameter estimation

a) Parameter estimation for the dynamic behavior

The parameter estimation is carried out with recursive least-squares estimation in the form of discrete square-root filtering (DSFI), see e.g. [4.4]. The basic model equation is

$$y(t) = \boldsymbol{\psi}^T(t)\hat{\boldsymbol{\theta}} + e(t) \tag{4.2.6}$$

and the data vector and the parameter estimation for the electrical part are

$$y(t) = U_A(t) \tag{4.2.7}$$

$$\boldsymbol{\psi}^T(t) = [I_A(t)\ \nu\omega_k(t)1] \tag{4.2.8}$$

$$\hat{\theta}^T = [\hat{\Theta}_1\ \hat{\Theta}_2\ \hat{\Theta}_3] \tag{4.2.9}$$

and for the mechanical part

$$y(t) = \dot{\omega}_k(t) \tag{4.2.10}$$

$$\boldsymbol{\psi}^T(t) = [I_A(t)\ \varphi_k(t)\ \omega_k(t)1] \tag{4.2.11}$$

$$\hat{\theta}^T = [\hat{\Theta}_4\ \hat{\Theta}_5\ \hat{\Theta}_6\ \hat{\Theta}_7] \tag{4.2.12}$$

Because of a fast input excitation the Coulomb-friction term is neglected and only the viscous-friction parameter M_{F1} is estimated under the condition that $|\omega_k| > 1.5$ rad/s.

The relation between the physical process coefficients and the parameter estimates are

$$\begin{aligned} &\hat{\Theta}_1 = R_A;\ \ \hat{\Theta}_2 = \Psi;\ \ \hat{\Theta}_3 = c_{0e} \\ &\hat{\Theta}_4 = \frac{\Psi}{\nu J};\ \ \hat{\Theta}_5 = -\frac{c_{S1}}{\nu^2 J};\ \ \hat{\Theta}_6 = -\frac{M_{F1}}{\nu^2 J};\ \ \hat{\Theta}_7 = -\frac{M_{S0}}{\nu^2 J} \end{aligned} \tag{4.2.13}$$

As the gear ratio ν is known, the ratio of inertia follows from

$$J = \frac{\hat{\Theta}_2}{\nu \hat{\Theta}_4} \tag{4.2.14}$$

All other process coefficients can be directly determined from the parameter estimates $\hat{\Theta}_i$.

For the parameter estimation the actuator operates in closed loop and the setpoint is changed with a PRBS between 10 and 70 deg. The derivatives $\omega_k = \dot{\varphi}_k$ and $\dot{\omega}_k = \ddot{\varphi}_k$ are determined by a state-variable filter with sampling time $T_{OSVF} = 2$ ms. The sampling time for the parameter estimation is $T_0 = 6$ ms. The resulting parameter estimates converge fast and the largest equation error is $\leq 5\%$ or ≤ 3.5 deg for the electrical part and $\leq 7 - -12\%$ for the mechanical part, [4.14]. Table 4.3 shows the deviations of the seven parameter estimates after introducing 14 different faults in several throttle actuators. All faults lead to different patterns, except F11 and F1, F2, which are not isolable.

b) Parameter estimation for the static behavior

In order to obtain more precise information on the mechanical part and especially the friction phenomena only the static behavior is considered for slow continuous input change according to a triangular upward and downward motion (phase 5 in Figure 4.10). Setting $\dot{\omega}_k = 0$ and neglecting the viscous friction (4.2.1) to (4.2.5) leads with $t = kT_0$ to

$$\begin{aligned} I_A(t) &= \frac{1}{\nu\Psi}\left(c_{S1}\varphi_k(k) + M_{S0} + M_{F0}\,\mathrm{sign}\,\omega_k(k)\right) \\ &= \boldsymbol{\Psi}^T(k)\boldsymbol{\Theta} \end{aligned}$$

Because of the direction-dependent Coulomb friction for the opening and closing two estimations are made:

Table 4.3. Process parameter deviations for different actuator faults: 0: no significant change; +: increase; −: decrease

		Features						
		parameter estimates						
	Faults	R_A	Ψ	c_{0e}	J	c_{S1}	M_{F1}	M_{S0}
F1	increase spring pre-tension	0	0	0	0	0	0	+
F2	decrease spring pre-tension	0	0	0	0	0	0	−
F3	commutator shortcut	−	−	0	+	+	+	0
F4	armature winding shortcut	0	−	0	+	+	+	0
F5	armature winding break	+	−	0	0	+	+	+
F6	additional serial resistance	+	0	0	0	0	0	0
F7	additional parallel resistance	−	−	0	0	+	+	0
F8	increased gear friction	0	0	0	+	+	+	0
F9	offset fault U_A	0	0	+/−	0	0	0	0
F10	offset fault I_A	0	0	−/+	0	0	0	+/−
F11	offset fault φ_K	0	0	0	0	0	0	−/+
F12	scale fault U_A	+/−	+/−	+/−	+/−	+/−	+/−	+/−
F13	scale fault I_A	−/+	0	0	+/−	+/−	+/−	+/−
F14	scale fault φ_K	0	−/+	0	−/+	−/+	−/+	−/+

$$\boldsymbol{\Psi}_1^T(k) = [\varphi_k^+(k) \;\; 1] \qquad \boldsymbol{\Psi}_2^T(k) = [\varphi_k^-(k) \;\; 1]$$

$$\hat{\boldsymbol{\Theta}}^+(k) = [\hat{\Theta}_1 \;\; \hat{\Theta}_2] \qquad \hat{\boldsymbol{\Theta}}^-(k) = [\hat{\Theta}_3 \;\; \hat{\Theta}_4]$$

with

$$\hat{\Theta}_1 = \frac{c_{S1}}{\nu\Psi} \qquad \hat{\Theta}_2 = \frac{M_{S0} + M_{F0}}{\nu\Psi}$$

$$\hat{\Theta}_3 = \frac{c_{S1}}{\nu\Psi} \qquad \hat{\Theta}_4 = \frac{M_{S0} - M_{F0}}{\nu\Psi}$$

The magnetic flux linkage Ψ is known from (4.2.13). The physical process parameters then result as

$$c_{S1} = \nu\Psi\frac{\hat{\Theta}_1 + \hat{\Theta}_3}{2}$$

$$M_{S0} = \nu\Psi\frac{\hat{\Theta}_2 + \hat{\Theta}_4}{2}$$

$$M_{F0} = \nu\Psi\frac{\hat{\Theta}_2 - \hat{\Theta}_4}{2}$$

The parameter estimation is performed with recursive DSFI and $T_0 = 6\,\text{ms}$ for each motion. Figure 4.12 shows the results for a fault-free case. The spring pre-tension M_{S0} leads to a positive offset of the linear spring characteristic and the dry friction shifts the friction characteristic by M_{F0}^+ and M_{F0}^- such that a hysteresis characteristic results. A comparison with the electrical torque $M_{el}' = \nu\Psi I_A$ related to the throttle axle indicates a good agreement with the estimated hysteresis characteristic. Changes

of the spring constant cause a change of the slope c_{S1} and changes of the friction or the pre-tension lead to shifts of the characteristics. (The oscillations of the calculated electrical torque are due to the closed-loop behavior in connection with adhesive friction or stick–slip effects, which are not modeled. The range around the point of returns, where adhesion works, is omitted in the parameter estimation for simplifying reasons).

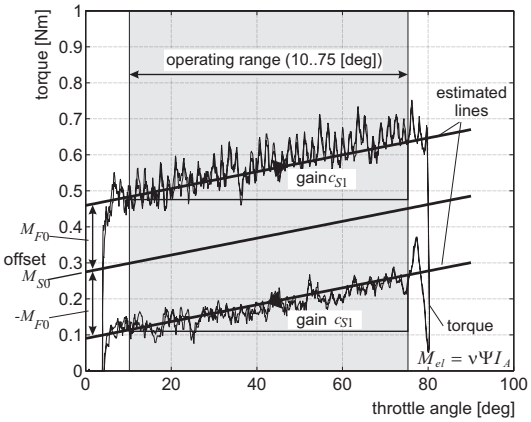


Fig. 4.12. Estimated static behavior of the throttle actuator (fault-free case)

4.2.4 Fault detection with parity equation

The application of several parity equations with structured residuals according to Section 3.1.2 was not successful, because the residuals have shown too much variance due to the difficulties in modeling the dynamic behavior of the mechanical system. This was especially due to the various kinds of dry and viscous friction and stick–slip effects of the motor brushes, the two-stage gear and pre-tensioned spring. Therefore only one parity equation is used for the electrical part. (4.2.1) leads to the voltage residual

$$r(t) = U_A(t) - R_A I_A(t) - \Psi \nu \omega_k(t) \tag{4.2.15}$$

The residual is calculated in continuous time. $\omega_k(t) = \dot{\varphi}_k(t)$ is determined by a state-variable filter (SVF). Because the SVF has to be applied to all measurements, all signals are low-pass filtered with an SVF with equal parameters. Simulations indicate that the residual shows a short burst for a sudden position sensor offset and an increase of the variance after changing the resistance R_A or the gain of the position sensor, after about 200 ms, [4.14].

However, the parameters R_A and Ψ change with temperature. The ranges are for -40 to $+120$°C, $R_A = 1.0 \ldots 1.81\,\Omega$ (rated value 1.3 Ω) and $\Psi = 0.0314 \ldots 0.0224$ Vs (rated value 0.028 Vs). This means that these parameters should

be taken from parameter estimation. An alternative is to apply adaptive parity equations, e.g. described in [4.3], [4.5] and [4.14]. The armature resistance changes follow from

$$\Delta \hat{R}_A = \frac{\sum_{i=0}^{N} \lambda^j I_A(k-i) r(k-i)}{\sum_{i=0}^{N} \lambda^j I_A(k-i) I_A(k-i)}$$

where $\lambda \leq 1$ is a forgetting factor. $\hat{R}_A$ is updated after a threshold ΔR_{Ath} is exceeded. As there exists a linear relationship between $\Psi(\vartheta)$ and $R_A(\vartheta)$ under the influence of the temperature ϑ, $\Psi(R_A)$ can be stored in a look-up table.

The theoretical and experimental investigations have shown that a detailed fault detection is especially possible with parameter-estimation methods. The parameter estimation of the static behavior gives best results for the mechanical part. Parameter estimation for the dynamic behavior allows one to detect faults in the mechanical part as well as the electrical part. Parity equations were only successful for the electrical part in combination with parameter estimation. The application of the parity equation for online real-time detection of position sensor faults and reconfiguration of the position control is described in [4.14]. Table 4.4 summarizes the conditions and applicability for the application of the fault-detection methods.

Table 4.4. Conditions and applicability for the application of fault-detection methods of the throttle actuator

Data evaluation	Parameter estimation dynamic	Parameter estimation static	Parity equation	Application
online real-time	dynamic excitation	not suitable	with parameter estimation	onboard
offline	dynamic excitation	rampwise test signals	with parameter estimation	quality control

4.2.5 Fault diagnosis

The experiments with the test cycle for quality control allows one to generate 30 different symptoms mainly by parameter estimation. In order to perform a fault diagnosis the observed symptoms can either be compared with a fault-symptom table, Table 4.3, or systematically evaluated with a fuzzy-logic diagnosis system, as described in [4.5]. The basis consists of fuzzy IF-THEN rules like

$$\text{IF}\,\{s_1 \text{ is } A_{11} \text{ AND } s_2 \text{ is } A_{21}\}\,\text{THEN}\,\{f \text{ is } F_1\}$$

Some examples are given in Figure 4.13. The membership functions of the attributes (symptoms) have simple shapes like triangles, trapezoids or ramps. Fuzzy-logic operators for the premise are min/max compositions. As it is sufficient to have one rule for each fault and to represent the faults as singletons $f_j \in 0 \ldots 1$, the degree of the fulfillment of the premise gives directly the degree of the conclusion, such that

more complicated evaluations of aggregation and defuzzification are avoided, see Figure 4.14. The diagnosis system was tested with different actuators and allowed to diagnose 38 different implemented faults (by measurement of only three signals).

In addition a learning neuro-fuzzy system SARAH (system for adaptive rule acquisition with Hebbian learning) [4.1] was applied with a classification rate close to 100% for 22 rules.

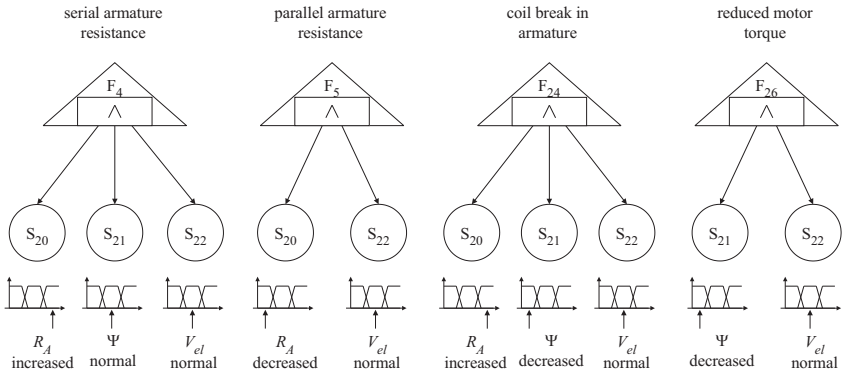


Fig. 4.13. Extract of fault-symptom relations with membership functions for the symptoms

4.2.6 Fault-diagnosis equipment

The developed test equipment consists as depicted in Figure 4.15 of a *personal computer* Pentium PC, 75 MHz), three boards for the real-time data processing (digital signal-processing (DSP) board dSpace DS 1003-192 with TMS 320C40 processor, 50 MHz, A/D-board with 32 channels, 16 bit, digital-I/O-board dSpace DS 4001, 32 TTL I/O, 5 timers 16 bit, 5 MHz) and the *actuator interface* (PWM power amplifier with MOSFET bridge, current limiter, signal conditioning for the three measurements U_A, I_A and φ_k and analog antialiasing low-pass filters).

The DSP board performs the real-time tasks for the control, the fault detection and diagnosis algorithms of the actuator. The algorithms are programmed with the language C. The operator interface is realized with the PC and the graphical user surface under Microsoft Windows. with the tool Testpoint 2.0 (Capital Equipment). The user can select automatic or manual test cycles and different displays and oscilloscope functions. Figure 4.16 shows an example. More details are given in [4.14] and [4.15].

The development of this diagnosis system was supported financially by the German Ministry of Education and Research (BMFT, 13 MV 01080) and managed by Forschungsvereinigung für Verbrennungskraftmaschinen (FVV), Frankfurt with project No. 540.

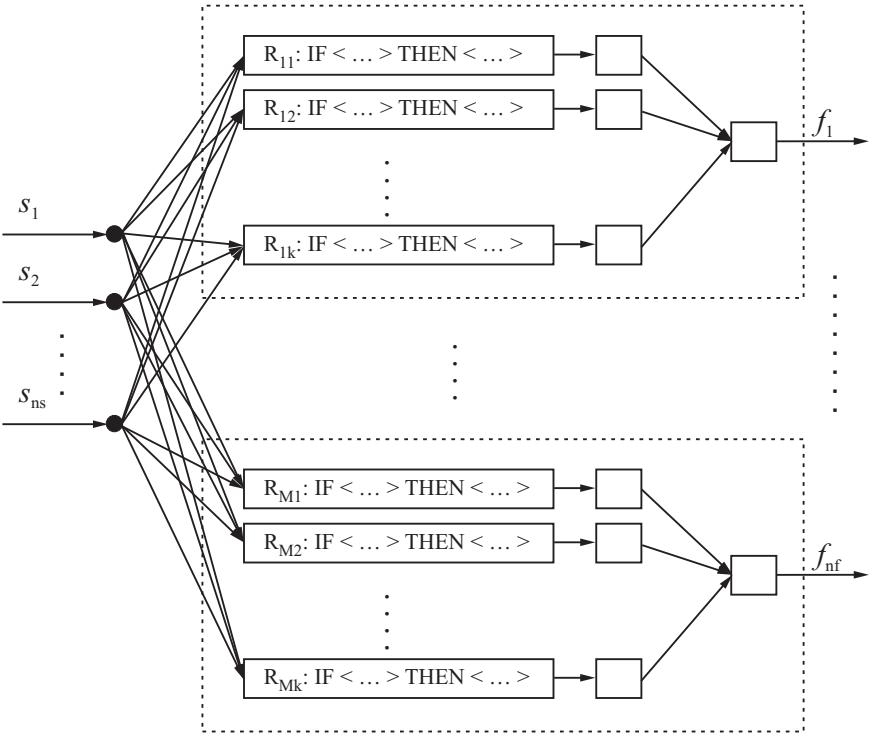


Fig. 4.14. Fuzzy-logic rule-based diagnosis

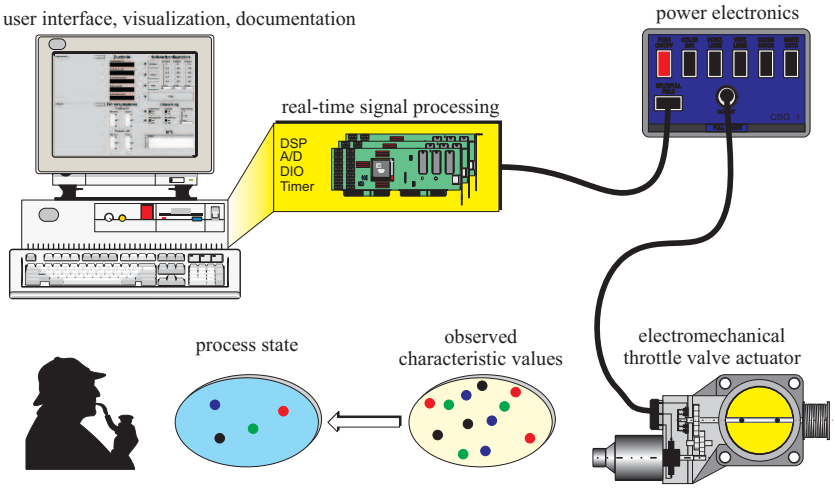


Fig. 4.15. Test equipment for the development of the throttle actuator fault diagnosis

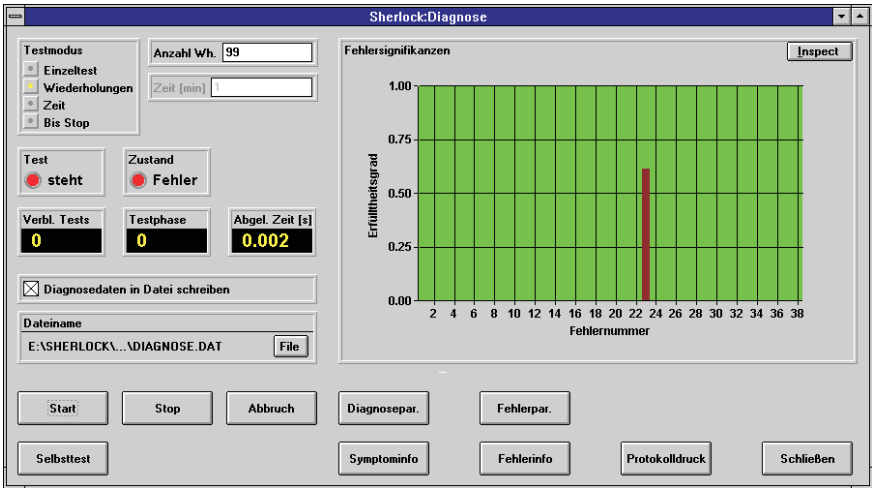


Fig. 4.16. Graphical user interface for the control and visualization of diagnosis tests

4.2.7 Conclusions

The developed fault detection and diagnosis methods for the throttle actuator can be applied as follows:

- quality control during manufacturing (e.g. end-of-line test) with special test cycles or for problematic returned actuators (offline data processing)
- function test and troubleshooting in service stations with built-in actuator (offline data processing)
- onboard fault detection during normal operation (online)
- onboard fault detection and control reconfiguration with redundant position sensors (fault tolerance).

For a comprehensive fault detection and diagnosis of the electrical and mechanical part mainly symptom generation with parameter estimation is suitable. Parity equations could only be applied for the electrical part.

4.3 Brushless DC motor and aircraft cabin pressure valve

4.3.1 Structure and models

The air pressure control in passenger aircraft is manipulated by DC motor driven outflow valves. The design of the outflow valve is made fault tolerant by two brushless DC motors which operate over the gear to a lever mechanism moving the flap, Figure 4.17.

The two DC motors form a duplex system with dynamic redundancy and cold standby, Figure 4.18. Therefore, a fault detection for both DC motors is required to switch from the possibly faulty one to the standby motor.

In the following it is shown how the fault detection was realized by combining parameter estimation and parity equations with implementation on a low-cost microcontroller, [4.9], [4.10], [4.11], [4.12].

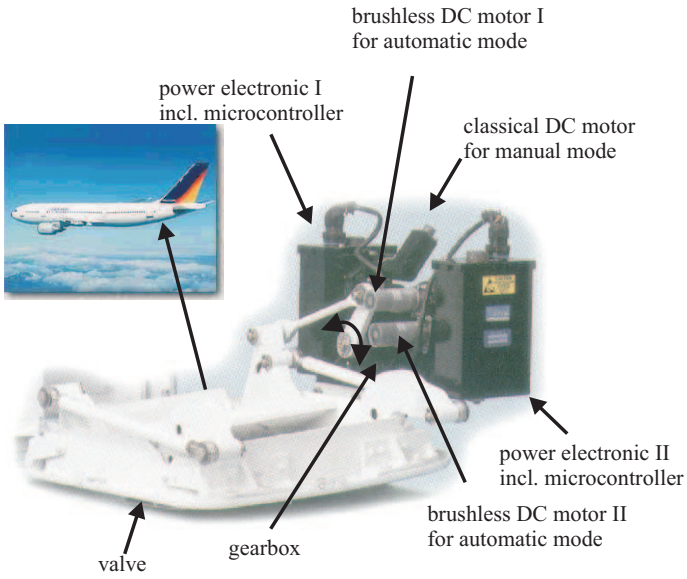


Fig. 4.17. Actuator servo-drive for cabin pressure control

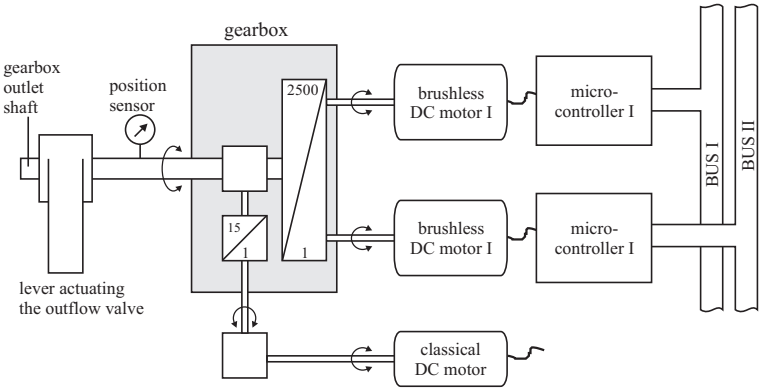


Fig. 4.18. Redundant DC motor drive system for the outflow valve

Figure 4.19 depicts the structure of the electronic commuted DC motor. The stator possesses three coils which are Y-connected and driven by a PWM (pulse-width modulation) inverter. The rotor has four permanent magnets. The position of the

rotor magnets is measured by three Hall sensors mounted on the stator. These determine the switching sequence of six MOSFET transistors of the PWM inverter. This switching scheme is implemented in a separated programmable logic array. The PWM inverter obtains a fixed voltage U_B and the resulting current I_B from the DC power supply and generates square wave voltages through the commutation logic via the six transistors to the three coils (phases).

The advantage of the electronic commutation is that no brushes exist, which are subject to wear and are a source of electromagnetic disturbances. Therefore, the reliability is relatively high. Possible faults in the brushless DC motor may originate from hall sensors, commutation circuits and transistors (overheating), stator coil windings, mechanical defects of bearings and magnets (eccentricity, striping), and electromagnetic disturbances. Usually, only measured signals for the supply voltage U_B, the input current I_B of the six-phase full bridge circuit and the angular rotor speed ω are available.

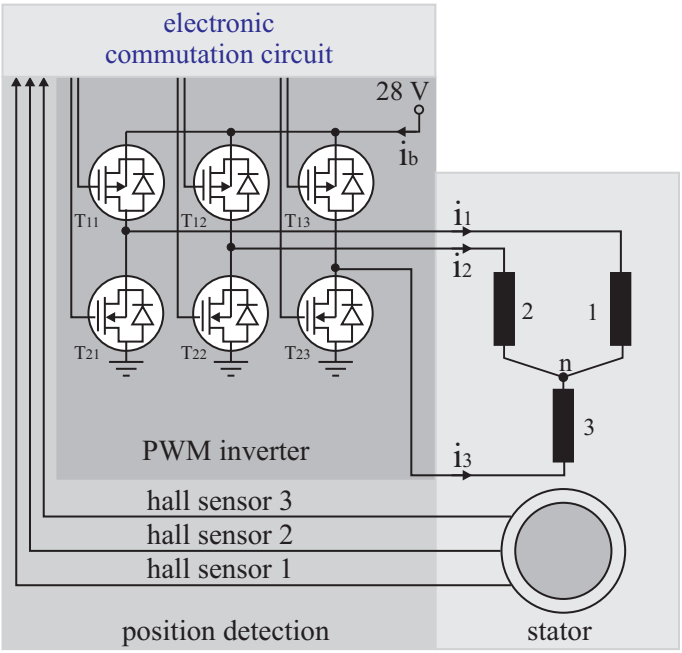


Fig. 4.19. Scheme of the brushless electronic commutated DC motor

A detailed model of the brushless DC motor for all three phases is given in [4.4] and [4.10]. It could be shown that for the case of fault detection averaged values (by low-pass filter) of the voltage $U(t)$ and the current $I(t)$ to the stator coils can be assumed. This leads to the voltage equation of the electrical subsystem

$$U(t) - k_E \omega_r(t) = RI(t) \tag{4.3.1}$$

with R the overall resistance and k_E the magnetic flux linkage. The generated rotor torque is proportional to the effective magnetic flux linkage $k_T < k_E$.

$$M_r(t) = k_T I(t) \tag{4.3.2}$$

(In ideal cases $k_E = k_T$.) The mechanical part is then described by

$$J_r \dot{\omega}_r(t) = k_T I(t) - M_f(t) - M_L(t) \tag{4.3.3}$$

with the ratio of inertia J_r, and the Coulomb friction torque

$$M_f(t) = c_f \operatorname{sign} \omega_r(t) \tag{4.3.4}$$

and the load torque $M_L(t)$. The gear ratio ν relates the motor shaft position φ_r to the flap position φ_g

$$\varphi_g = \varphi_r / \nu \tag{4.3.5}$$

with $\nu = 2500$. The load torque of the flap is a normal function of the position φ_g

$$M_L = c_s f(\varphi_g) \tag{4.3.6}$$

and is approximately known around the steady-state operation point. (For the experiments the flap was replaced by a lever with a spring.)

4.3.2 Fault detection with parameter estimation

For fault detection the following measurements are available: $U(t)$, $I(t)$, $\omega_r(t)$, $\varphi_g(t)$. Using the notation

$$y(t) = \boldsymbol{\psi}^T(t)\boldsymbol{\theta} \tag{4.3.7}$$

two equations were used for *parameter estimation*:

- electrical subsystem,

$$y(t) = U(t), \Psi^T(t) = [I(t)\, \omega_r(t)]; \quad \theta^T = [R\, k_E] \tag{4.3.8}$$

- mechanical subsystem,

$$\begin{aligned} y(t) &= k_T I(t) - c_s f(\varphi_g(t) - J_r \dot{\omega}_r(t)) \\ \Psi^T(t) &= [\operatorname{sign} \omega_r(t)];\ \theta^T = [c_f] \ \ (J_r \text{ known}) \end{aligned} \tag{4.3.9}$$

Hence, three parameters $\hat{R}$, $\hat{k}_E$ and $\hat{c}_f$ are estimated. Various parameter estimation methods were applied like: RLS (recursive least squares), DSFI (discrete square-root filtering), FSDFI (fast DSFI), NLMS (normalized least mean squares) and compared with regard to computational effort in floating-point and integer word realization and estimation performance. The floating-point implementation is standard for, e.g. 16-bit signal processors and in this case RLS, DSFI or FDFSI can be used. However, integer word implementation is (still) required if reliable and certified low-cost microcontrollers like the 16-bit Siemens C 167 have to be used. Then only NLMS is feasible, [4.9].

4.3.3 Fault detection with parity equations

The *parity equations* are obtained from the basic two equations (4.3.1) and (4.3.3) by assuming known parameters (obtained from parameter estimation)

$$r_1(t) = U(t) - RI(t) - k_E\omega_r(t) \tag{4.3.10}$$

$$r_2(t) = k_T I(t) - J_r\dot{\omega}_r(t) - c_f \operatorname{sign}\omega_r(t) - c_s f(\varphi_g) \tag{4.3.11}$$

$$r_3(t) = U(t) - \frac{R}{k_T}(J_r\dot{\omega}_r(t) + c_s f(\varphi_g) + c_f \operatorname{sign}\omega_r(t) + k_E\omega r(t)) \tag{4.3.12}$$

$$r_4(t) = \varphi_g(t) - \varphi_r(t)/\nu \tag{4.3.13}$$

Each of the residuals is decoupled from one measured signal. r_1 is independent from φ_g, r_2 from U, r_3 from I, r_4 from all but φ_r. (φ_r is assumed to be correct. It can directly be supervised by a logic evaluation within the motor electronics.) Figure 4.20 shows measured signals, parameter estimates and residuals for five different implemented faults.

The actuator was operating in closed loop with slow triangle changes of the reference variable (setpoint). The fault-detection methods, including state-variable filters (SVF) were implemented on a digital signal processor TI TMS 320 C40 with signal sampling period $T_0 = 1$ ms. The results for fault detection are summarized in Table 4.5. The sign and size of changes for the parameter estimates with FDSFI clearly allow one to identify the parametric faults and for the parity residuals the respective additive (offset) sensor faults. But there are also cross-couplings: for parametric faults some residuals show changes and for sensor-additive faults some parameter estimates change (except for φ_g), which can all be interpreted by the equations used. According to [4.2] the symptom pattern is weakly isolating as a parametric fault of R and an additive fault in U differ only in one symptom. However, all faults can be isolated. Including the standard deviation of the symptoms isolability can be improved, [4.9]. By processing eight symptoms with a rule-based fuzzy-logic diagnosis system, finally 10 different faults could be diagnosed, [4.9], [4.13].

Table 4.5. Parameter deviations and parity equation residuals for different actuator faults (0 no significant change; + increase; ++ large increase; − decrease; −− large decrease)

	Parameter estimates			Residual parity equations			
Faults	$\hat{R}$	$\hat{k}_E$	$\hat{c}_f$	r_1	r_2	r_3	r_4
increasing R	+	0	0	+	0	+	0
increasing c_f	0	0	++	0	−−	++	0
offset U	+	+	0	++	0	++	0
offset φ_g	0	0	0	0	0	0	0
offset I_b	++	−	−−	++	++	0	0

Because the position sensors of the rotor φ_r and the shaft φ_g yield redundant information, sensor fault detection for φ_g was used to reconfigure the closed loop

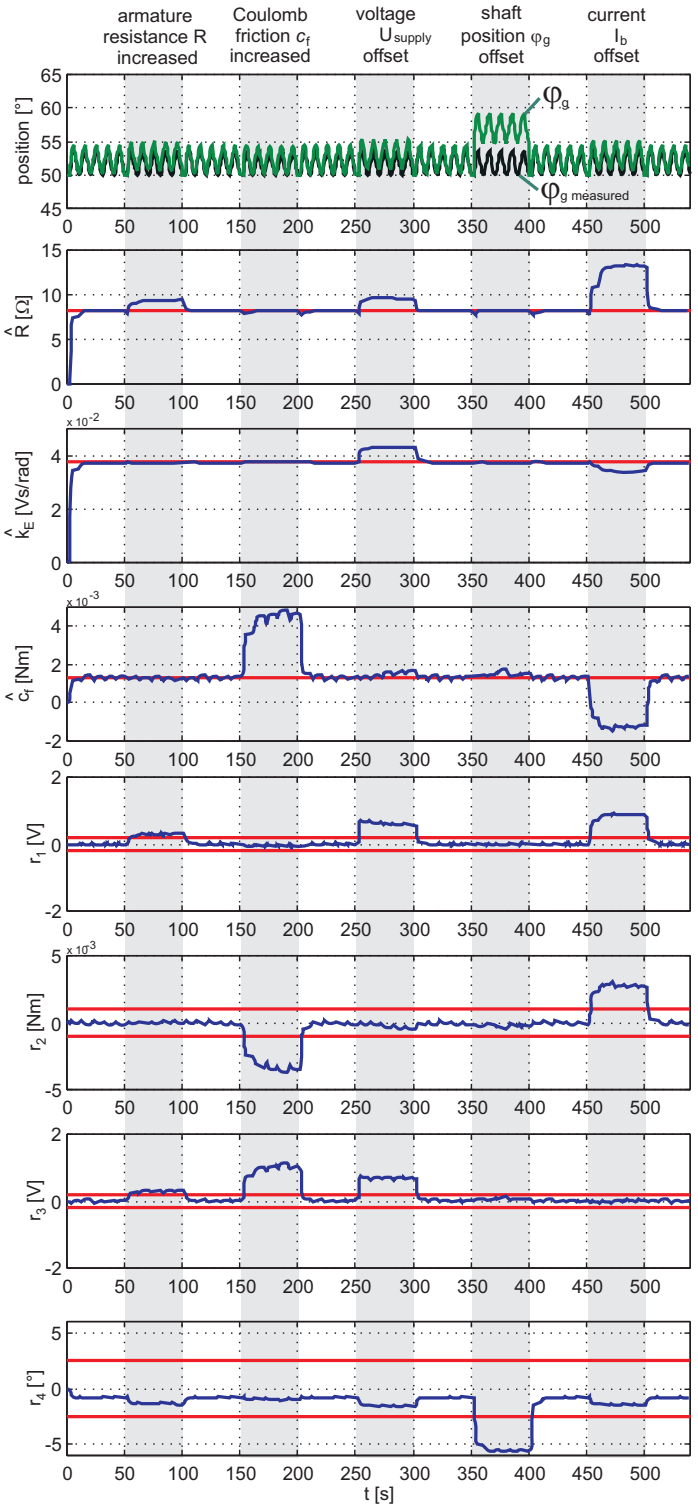


Fig. 4.20. Resulting symptoms from parameter estimation and parity equations by measuring $U(t)$, $I(t)$, $\omega(t)$, $\varphi_g(t)$ and $\varphi_r(t)$

after failure of φ_g by using φ_r as the control variable, [4.9]. The described combined fault-detection methodology needs about 8 ms calculation time on a 16-bit microcontroller. Therefore, online implementation in a smart actuator is possible by only measuring four easy accessible variables U, I and ω_r and φ_g.

4.3.4 Conclusions

The fault-diagnosis approach of this actuator has shown that the brushless DC motor can be modeled as for DC motors with brushes. Parameter estimation is primarily suitable to detect parametric faults, and parity equations to detect additive faults. If the input signal U stays approximately constant, only parity equations should be applied, which then may indicate faults. Then for isolating or diagnosing the faults a test signal on U can be applied for short time to gain deeper information. Hence, by applying both parameter estimation and parity equations a good fault coverage can be obtained.

5

Fault diagnosis of fluidic actuators

Fluidic actuators are characterized by their rugged design and high power-to-weight ratio. They enable one to generate linear motions easily and directly by employing hydraulic cylinders and diaphragm drives. They also dissipate energy only if in dynamic operation and during holding phases generate high reaction forces with little energy consumption. Characteristics and mathematical models are treated, e.g. in [5.9], [5.13], [5.14], [5.20] and [5.21]. Methods for fault detection of two basic actuator principles, a hydraulic servo cylinder and a pneumatic diaphragm flow valve are considered in the sequel.

5.1 Hydraulic servo axis[1]

Fault detection of hydraulic systems is classically mainly based on the supervision of *directly measurable quantities*, e.g. differential pressure across a filter element to detect clogging or the monitoring of hydraulic fluid for debris, [5.29]. In order to gain insight into the process, additional special sensors can be provided. However, they increase cost and are themselves susceptible to faults. An overview of advanced methods including oil condition monitoring is given by [5.21].

[5.12] developed fault-detection methods for a linear hydraulic servo-axis based on frequency response measurements as well as by the application of the extended Kalman filter (EKF). Application of parameter estimation is reported by [5.26]. Based on a linearized valve flow curve, [5.18] also utilized the EKF to extract physical quantities (Coulomb friction, the amount of throttling by the compensation conduits and pressure-induced forces) from the system behavior of a fault-tolerant proportional valve, see also [5.1]. [5.27] investigated the transfer of fault-detection methods to a microcontroller. Physical quantities from the process such as the coefficient of internal leakage or the degree of control edge erosion were estimated by means of an EKF. Experimental results are shown for linear models and for constant fluid temperature. The fault-detection methods that include the EKF typically employ the

[1] Worked out by Marco Münchhof.

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_5, © Springer-Verlag Berlin Heidelberg 2011

valve spool displacement sensor, chamber pressure sensors and the piston displacement sensor as well as information about the supply pressure. However, a precise description of hydraulic systems requires nonlinear models, see also [5.14].

The following section describes the methods and experimental results for a relatively comprehensive fault detection and diagnosis procedure and use of standard sensors.

5.1.1 Hydraulic servo axis structure

Figure 5.1 shows a scheme of a typical hydraulic linear actuator with power supply. The hydraulic power supply consists of an oil tank and electrically driven piston pump connected to a proportional-acting electro-hydraulic servovalve or spool valve. This servo-valve is position controlled. Depending on its position, the cylinder moves with a corresponding direction and speed. A return valve avoids backflow and a pressure-relief valve acts as a safeguard against excessive pressures. A hydraulic accumulator damps the pressure oscillations from the piston pump. A view of the testrig used for the experiments is shown in Figure 5.2.

The pressure supply consists of a swash plate pump, a proportional valve, and a differential cylinder with integrated displacement sensor. The control scheme and the measurements used are depicted in Figure 5.3. The position of the piston (travel way) is the main control variable. Its position controller is a digital controller which has been programmed as a simple P-controller and acts on the reference variable of the proportional valve position controller.

Standard measurements are the pump pressure p_P, the valve spool position y_V and the piston position y. The chamber pressure p_A and p_B are additionally used for the experiments, but are in general optional for servo axes.

5.1.2 Faults of hydraulic servo axes

A list of typical faults for hydraulic servo axes is given in Table 5.1. The *hydraulic fluid* may contain *gas enclosures*, such as undissolved air, vapor and foam. Undissolved air is typically the result of insufficient bleeding, which is one of the most common faults during commissioning of hydraulic systems. Vapor enclosures can be the result of inadequate operating conditions. Foam results if the hydraulic storage tank is designed in an inappropriate way. The *proportional valve* may show *grooving* and *erosion*. Grooving stems from splinters being caught between the valve spool and the housing. Upon moving of the valve spool, these splinters start to grind themselves into the valve spool. Erosion of the control edges can be caused by cavitation and jet erosion. The latter is caused by small particles being suspended in the hydraulic fluid. The *hydraulic cylinder* may suffer from *leakages* and *gas enclosures*. The two cylinder chambers are separated by the piston and by sealing. This sealing may wear out over time and allows more and more fluid to leak from the higher pressurized to the lower pressurized chamber (*internal leakages*). Similarly, oil can leak from the sealing between the pushrod and the cylinder housing.

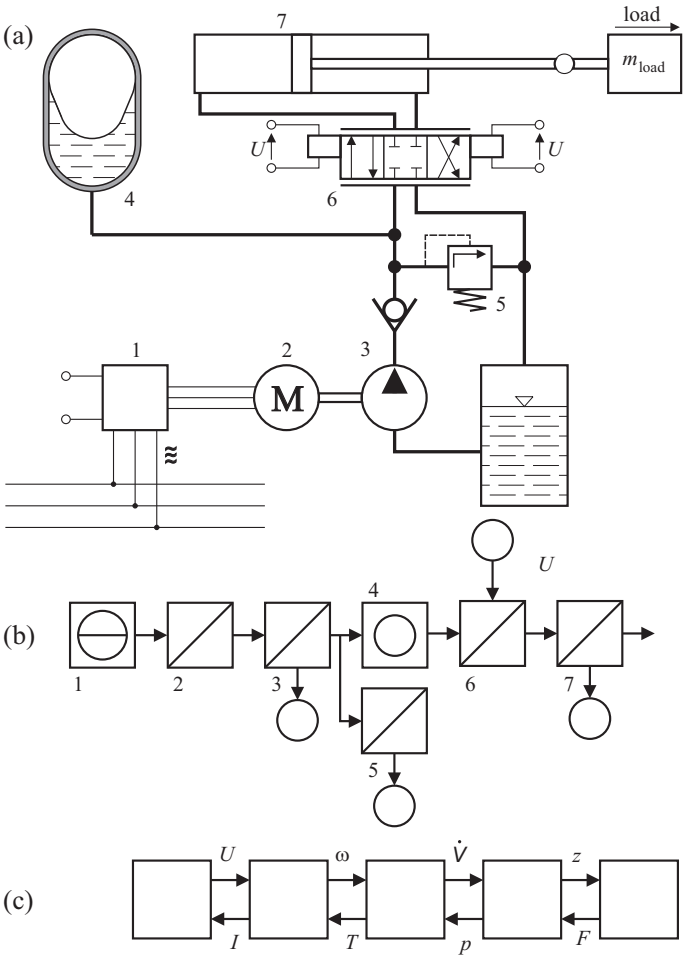


Fig. 5.1. Scheme of a hydraulic servo axis with power supply: a) schematic, b) energy flow scheme, c) two-port representation. 1 power electronics, 2 AC motor, 3 axial piston pump, 4 accumulator, 5 check valve, 6 proportional valve, 7 cylinder

Furthermore, the sensors mounted at the hydraulic servo axis may be faulty. Sensor faults can have manifold reasons and are typically not characterized in detail. Rather, they are classified by their effect on the sensor signal as additive and multiplicative sensor faults and total loss of the measurement.

It is difficult to obtain failure rates of hydraulic systems, because they are usually not published. An exemption is the area of aircraft, [5.20], according to refurbishments for airlines. The component statistics are divided into lap assemblies, valve assembly, cylinders, mechanical defects, power control unit and others. Tables 5.2 and 5.3 show the obtained failure rates based on about 3000 parts, see also Figures 5.4 and 5.5.

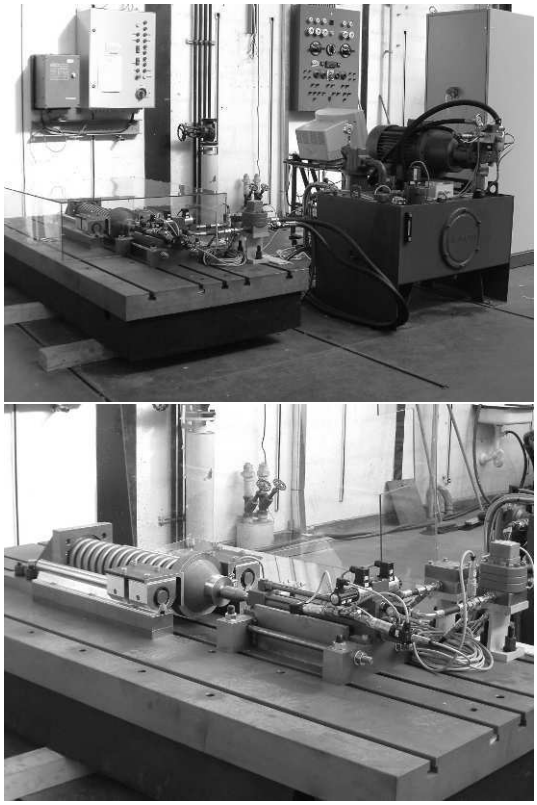


Fig. 5.2. The total testrig (top); the linear servo axis in detail (bottom)

Table 5.1. Typical faults of hydraulic components

Subsystem	Fault
Hydraulic fluid	Gas enclosures
Proportional valve	Grooving
	Erosion of control edges
	Increased friction
	Change of or wrong fit
	Break of winding
Cylinder	Internal leakages
	External leakages
	Gas enclosures
	Increased friction
	Slack between cylinder and external load
Connection lines	Leakage
Accumulator	Loss of charge pressure

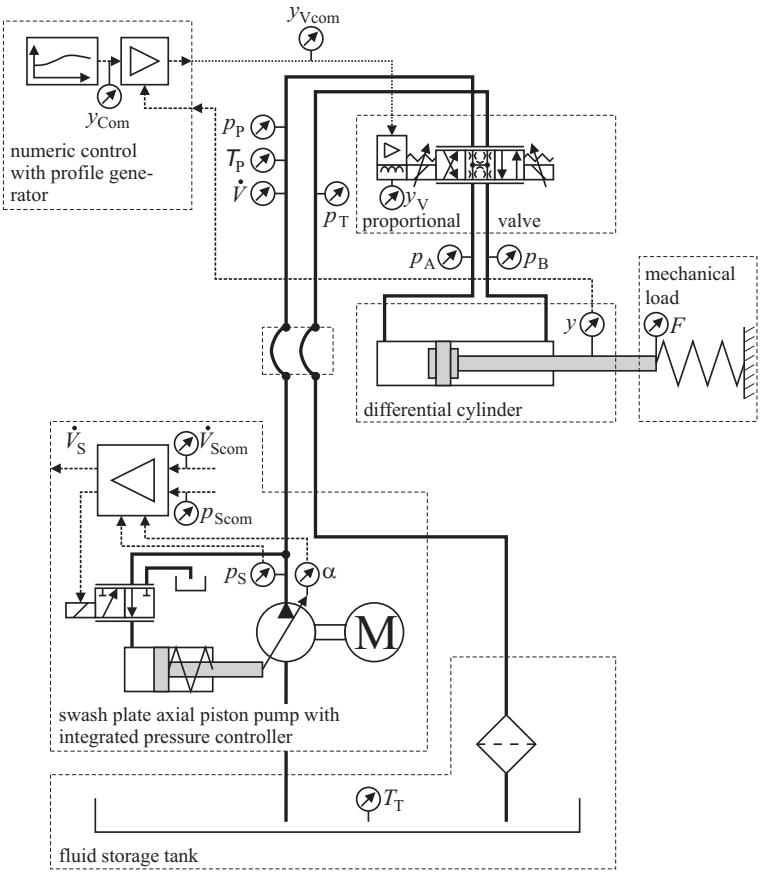


Fig. 5.3. Control scheme and measurement of the servo axis; p_P pump pressure, p_A pressure chamber A, p_B pressure chamber B, y_V valve spool position, y piston position, T_p oil temperature before spool valve, p_{Scom} pressure setpoint and $\dot{V}_{Scom}$ volume flow rate setpoint for integrated pump controller

Table 5.2. Failure rates of servo-hydraulic components (approx. 3000 parts), [5.20]

Lap failures	Spool ass'y	Valve ass'y	Cylinder	Mechanics	Power control unit	Other
%	32	19	16	14	3	14

Table 5.3. Failure rates of laps, valves and cylinder, [5.20] (estimates based on long-term experience)

Lap failures	Erosion	Damage due to debris	External leakage	Other
%	65	20	10	5

Valve failures	Internal leakage	Improper function	External leakage	Other
%	35	35	15	15

Cylinder failures	External leakage	Internal leakage	Broken or cracked	Other
%	58	14	14	14

About 50% of returned components are caused by faults in the lap and valve assembly. Therefore hydraulic valves are the most susceptible components, [5.20]. Most faults in the lap assembly, consisting of a slide and a sleeve are due to erosion or damage because of debris. Faults in the valve assembly are mainly internal and external leaks or unspecifically, improper function. The cylinders are not so often faulty as the valves. The main reasons are external and internal leakages (seals) or breaks and cracks, mainly caused by wear.

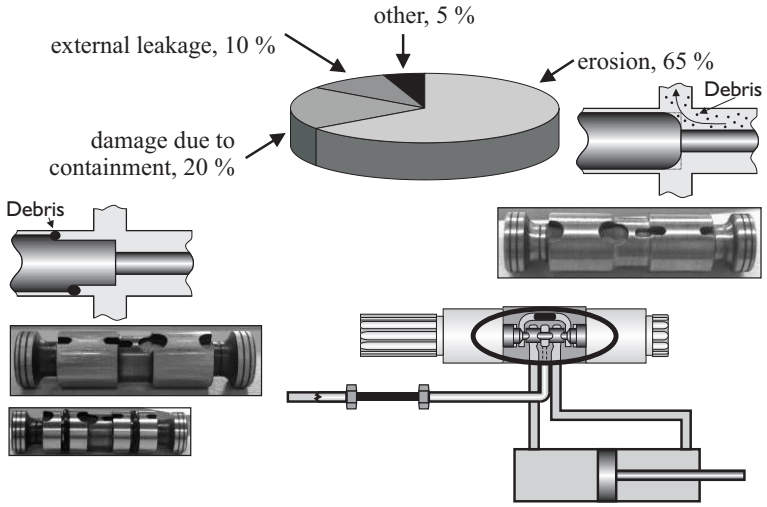


Fig. 5.4. Possible faults in proportional valves

Figure 5.6 depicts the measurements used and different steps to the fault diagnosis of a hydraulic servo axis as developed by [5.20].

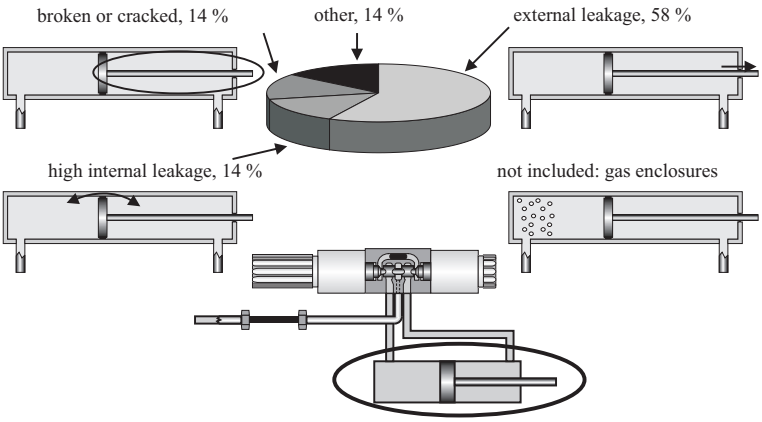


Fig. 5.5. Possible faults in hydraulic cylinders

5.1.3 Models of spool valve and cylinder

a) Component specific models

Inside the valve, the fluid flows through the *valve orifices*. The *control edges* are shaped such that a turbulent flow evolves, which in general should not be temperature dependent. However, measurements at the testbed have shown a temperature dependency which can most likely be attributed to changes in the fluid density ρ. The flow over a control edge (see Figure 5.7b) is given as

$$\dot{V} = \alpha_{\mathrm{D}}(y_V) A(y_V) \sqrt{\frac{2}{\rho(p,T)}} \sqrt{|\Delta p(t)|}\, \mathrm{sign}\, \Delta p(t) \tag{5.1.1}$$

where all temperature and displacement dependent terms have been collected into the term b_{V}, thus

$$\dot{V}(\Delta p, T, y_V) = b_{\mathrm{V}}(y_V, T) \sqrt{|\Delta p(t)|}\, \mathrm{sign}\, \Delta p(t) \tag{5.1.2}$$

with

$$b_{\mathrm{V}}(y_V, T) = \alpha_{\mathrm{D}}(y_V) A(y_V) \sqrt{\frac{2}{\rho(p,T)}} \tag{5.1.3}$$

b) Cylinder model

Based on the mass balance, the pressure buildup in the hydraulic chambers (here chamber A, see Figure 5.7a) is governed by

$$(V_{oA} + A_A\, y(t)) \frac{1}{E(p_A, T)} \dot{p}_A + A_A \dot{y}(t) = \left(\dot{V}_A(t) - \dot{V}_{AB}(t) \right) \tag{5.1.4}$$

where E is the bulk modulus. The term V_{oA} denotes the oil flow from the valve. Other volume flows present in hydraulic cylinders are the leakage flows. Two leakage flows can be differentiated in a double-acting cylinder, one between chamber A and B and one between chamber B and the surroundings. These two flows are modeled as laminar flows, thus, e.g. for the flow between chambers A and B,

$$\dot{V}_{AB}(p_A, p_B, T) = G_{\mathrm{AB}}(T)\,(p_A(t) - p_B(t)) \tag{5.1.5}$$

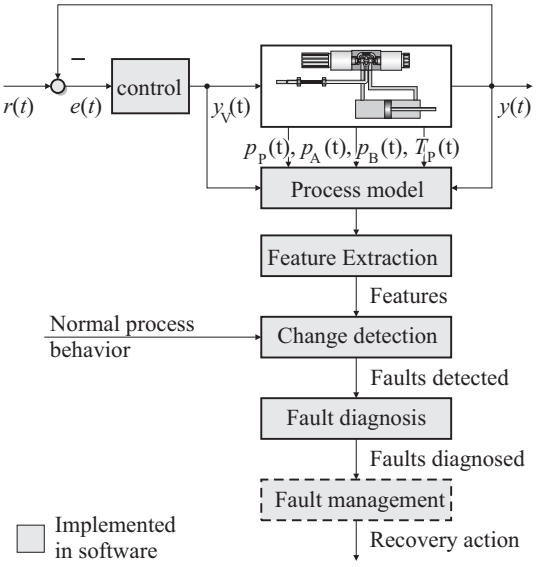


Fig. 5.6. Scheme of fault detection and diagnosis

c) Combined model and choice of model output

These component models, (5.1.2), (5.1.4), and (5.1.5) can now be combined into one differential equation:

$$(V_{oA} + A_A y(t))\,\frac{1}{\bar{E}(T)}\,\dot{p}_A(t) + A_A \dot{y}(t) = \dot{V}_A(p_A, p_P, T, y_V) - G_{\mathrm{AB}}(T)\,(p_A(t) - p_B(t)) \tag{5.1.6}$$

with the flow from the valve to the cylinder chamber A being given as

$$\dot{V}_A(p_A, p_P, T, y_V) = \begin{cases} b_{\mathrm{V2}}(y_V, T)\sqrt{|p_P(t) - p_A(t)|}\,\mathrm{sign}\,(p_P(t) - p_A(t)) & \text{for } y_V > 0 \\ b_{\mathrm{V1}}(y_V, T)\sqrt{|p_A(t)|}\,\mathrm{sign}\,(p_A(t)) & \text{for } y_V < 0 \end{cases} \tag{5.1.7}$$

As a simplification, the pressure dependence of the bulk modulus has been neglected and an average bulk modulus is assumed. Furthermore, the pressure at port T (tank)

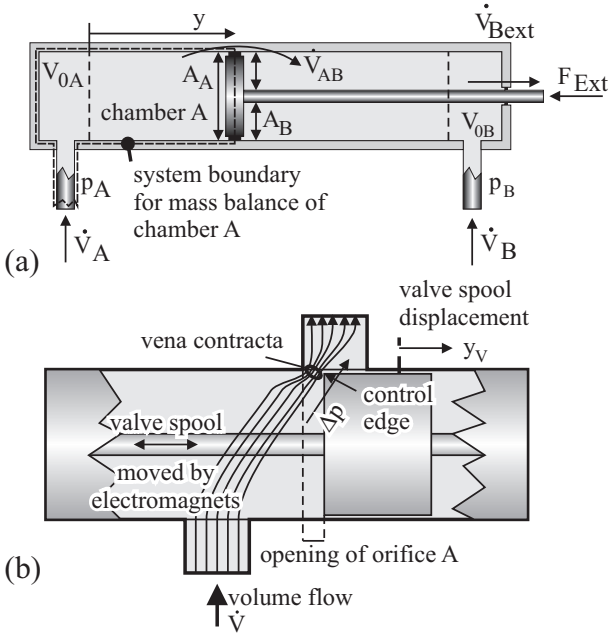


Fig. 5.7. Scheme of hydraulic cylinder (a) and valve orifice (b)

is set to zero. The differential equation (5.1.6) contains two derivatives with respect to time: The derivative of the chamber pressure $p_A(t)$ and the derivative of the piston displacement $y(t)$. So far, in [5.12], [5.18], [5.27], (5.1.6) was integrated in variable $p_A(t)$ with respect to time, resulting in the ordinary differential equation (ODE)

$$\dot{p}_A(t) = \frac{\bar{E}(T)\left(\dot{V}_A(p_A, p_P, T, y_V) - G_{\mathrm{AB}}(T)\,(p_A(t) - p_B(t)) - A_A\dot{y}(t)\right)}{V_{oA} + A_A y(t)} \tag{5.1.8}$$

(5.1.6) can also be integrated with respect to the piston displacement, yielding the ODE

$$\dot{y}(t) = \frac{1}{A_A}\left(\dot{V}_A(p_A, p_P, T, y_V) - (V_{oA} + A_A y(t))\frac{1}{\bar{E}(T)}\dot{p}_A(t)\right) \tag{5.1.9}$$

As shown in [5.20], (5.1.8) is very sensitive to measurement noise and shows a lot of oscillatory behavior. Therefore (5.1.9) is used as a basis for the formulation of a parity equation, since even this white-box model, based exclusively on physical laws and parameterized with values taken from data sheets, already shows a very good model fidelity. Based on an investigation of different sample times, experiments have shown that the model can be calculated with a sample time of $T_0 = 2\,\mathrm{ms}$.

c) Parameterization of valve curve

At the testbed, a proportional valve 4WREE with a V1 spool shows a nonlinear relation between the valve spool position and the resulting valve opening, as can be seen in Figure 5.8. This is typical for many proportional-acting direct-driven valves, since the special shaping of the flow characteristics helps in achieving better control performance. A linear approximation of valve flow characteristics can hardly capture such a nonlinear behavior, as was shown in [5.20]. Therefore, a polynomial approximation of the coefficient of valve flow is used as

$$\dot{V}_A(p_A,\, p_P,\, T,\, y_V) = \begin{cases} \sum_{i=k}^{l} \left(b_{1i}(T) y_V(t)^i\right) \sqrt{|p_P(t) - p_A(t)|}\, \text{sign}\,(p_P(t) - p_A(t)) \\ \text{for } y_V \geq 0 \\ \sum_{i=k}^{l} \left(b_{2i}(T) y_V(t)^i\right) \sqrt{|p_A(t)|}\, \text{sign}\,(p_A(t)) \\ \text{for } y_V < 0 \end{cases} \tag{5.1.10}$$

A very good approximation has been found for $k = 0$ and $l = 4$, as can be seen in Figure 5.8. An alternative is the set of $k = 1$ and $l = 3$, which limits the computational impact even more at the price of a slight increase in model fidelity. The resulting model fidelity is illustrated in Figure 5.9. This model was parameterized with a dataset recorded at a fluid temperature of 20° C.

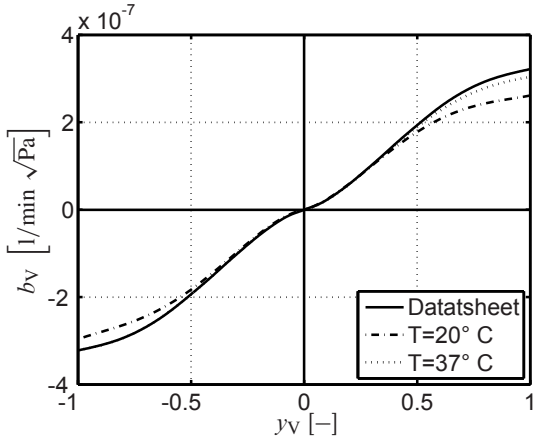


Fig. 5.8. Polynomial approximation of valve flow characteristics

d) Temperature dependency

As was witnessed in Figure 5.8, the characteristics of the valve flow change with temperature. Experiments by [5.20] have shown that in a range from $T_{P1} = 20°\,\text{C}$ to $T_{P2} = 40°\,\text{C}$, the model parameters change almost linearly with temperature. Hence, a linear approximation is applied for a temperature adaptation of the model,

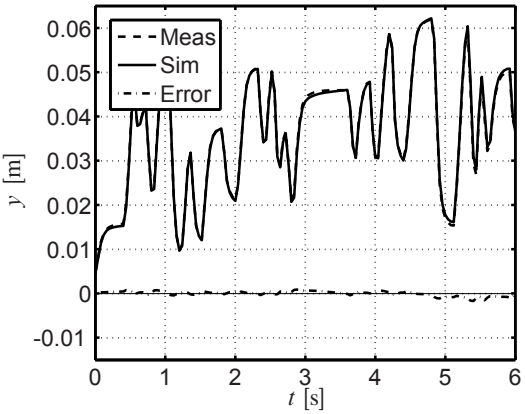


Fig. 5.9. Comparison of measured and modeled piston displacement y with polynomial approximation of b_V

$$b_{ij}(T_P) = \frac{b_{ij}(T_{P2}) - b_{ij}(T_{P1})}{T_{P2} - T_{P1}} \; (T_P - T_{P1}) + b_{ij}(T_{P1}) \tag{5.1.11}$$

where the model parameters have been determined a priori for the temperatures $T_{P1} = 20°\,\mathrm{C}$ and $T_{P2} = 40°\,\mathrm{C}$. With these suggested improvements, the model yields a fidelity that is suitable for detection of tiny faults in online fault detection and also for analytical redundancy at *changing fluid temperatures*. It is obvious that such a precise model can also be used as a "model sensor" in the case of a sensor fault in y.

5.1.4 Fault detection and diagnosis of valve and cylinder

Based on the described models, a fault detection and diagnosis system has been derived as shown in Figure 5.10. The model for valve and cylinder can be extended in a straightforward way to chamber B.

a) Parity equations

By using various subsets of sensors mounted at the testbed, a total of six parity equations can be formulated,

$$r_1(t) = y(t) - \hat{y}_1(p_\mathrm{P}, p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.12}$$
$$r_2(t) = y(t) - \hat{y}_2(p_\mathrm{S}, p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.13}$$
$$r_3(t) = y(t) - \hat{y}_3(p_\mathrm{P}, p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.14}$$
$$r_4(t) = y(t) - \hat{y}_4(p_\mathrm{S}, p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.15}$$
$$r_5(t) = y(t) - \hat{y}_5(\dot{V}, p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.16}$$
$$r_6(t) = y(t) - \hat{y}_6(p_\mathrm{A}, p_\mathrm{B}, T_\mathrm{P}, y_\mathrm{V}) \tag{5.1.17}$$

where $r_1(t)$ and $r_3(t)$ are formulated based on the mass balance of chamber A and $r_2(t)$ as well as $r_4(t)$ are based on the mass balance for chamber B. Residual $r_5(t)$ is the only residual driven by a flow meter which is not a standard sensor. Taking into account the large costs of a flow meter and its limited performance, it is not advisable to use it for fault detection and diagnosis. Figure 5.11a) shows the behavior of the residual $r_1(t)$ in the fault-free case and Figure 5.11b) illustrates the reaction of the residual to control edge erosion for a family of experiments. The influence of sensor faults is shown in Figure 5.12, where the valve spool displacement sensor signal has an offset of $\Delta y_\mathrm{V} = 0.002\,\mathrm{m}$.

b) Parameter Estimation

The precise models can also be used as a basis for an LS parameter estimation approach, which extracts the coefficients of valve flow, the bulk modulus and the coefficient of laminar leakage flow from a set of measurements. Since parameter estimation is based on sampled signals, the time t will now be expressed as integer multiples k of a fixed sampling time T_0. Based on the model (5.1.9) and (5.1.10) derived in the previous sections, a parameter estimation problem can be formulated.

This parameter estimation problem will be split up into the data matrix where kT_0 is abbreviated by k

$$\vec{\Psi}^T = \begin{pmatrix} \sqrt{|p_\mathrm{P}(k) - p_\mathrm{A}(k)|}\mathrm{sign}\,(p_\mathrm{P}(k) - p_\mathrm{A}(k)) \cdot (y_\mathrm{V}(k) \geq 0) & \dots \\ y_\mathrm{V}(k)\sqrt{|p_\mathrm{P}(k) - p_\mathrm{A}(k)|}\mathrm{sign}\,(p_\mathrm{P}(k) - p_\mathrm{A}(k)) \cdot (y_\mathrm{V}(k) \geq 0) & \dots \\ \vdots & \\ y_\mathrm{V}(k)^4\sqrt{|p_\mathrm{P}(k) - p_\mathrm{A}(k)|}\mathrm{sign}\,(p_\mathrm{P}(k) - p_\mathrm{A}(k)) \cdot (y_\mathrm{V}(k) \geq 0) & \dots \\ \sqrt{|p_\mathrm{A}(k)|}\mathrm{sign}\,(p_\mathrm{A}(t)) \cdot (y_\mathrm{V}(k) < 0) & \dots \\ y_\mathrm{V}(k)\sqrt{|p_\mathrm{A}(k)|}\mathrm{sign}\,(p_\mathrm{A}(k)) \cdot (y_\mathrm{V}(k) < 0) & \dots \\ \vdots & \\ y_\mathrm{V}(k)^4\sqrt{|p_\mathrm{A}(k)|}\mathrm{sign}(p_\mathrm{A}(k)) \cdot (y_\mathrm{V}(k) < 0) & \dots \\ (p_\mathrm{A}(k) - p_\mathrm{B}(k)) & \dots \\ A_\mathrm{A}\; y(k)\; \dot{p}_\mathrm{A}(k) & \dots \end{pmatrix} \tag{5.1.18}$$

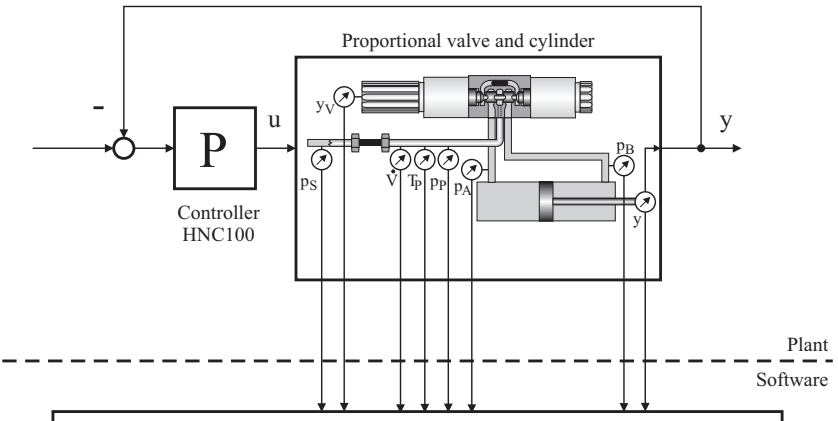
where $(y_\mathrm{V}(k) \geq 0)$ is one if the condition is met and zero otherwise. The output vector is written as

$$\vec{y}^T = (A_\mathrm{A}\dot{y}(k), \dots) \tag{5.1.19}$$

The solution of

$$\hat{\vec{\theta}} = \left(\vec{\Psi}^T\vec{\Psi}\right)^{-1}\vec{\Psi}^T\,\vec{y} \tag{5.1.20}$$

supplies estimates for the parameter vector $\vec{\theta}$ given as



Fault diagnosis

		r1	r2	r3	r4	(r5)	r6
Pressure Supply	Pressure Discharge Valve (F1)	0	0	0	0	0	0
	Supply Line Congestion (F2+4)	-	-	+	+	0	0
	Return Line Congestion (F3+5)	+	+	-	-	0	+
Valve	Eroded Control Edge A-T (F6.1)	-	-	0	0	0	0
	Eroded Control Edge P-A (F6.2)	-	-	0	0	0	0
	Eroded Control Edge P-B (F6.3)	0	0	+	+	0	0
	Eroded Control Edge B-T (F6.4)	0	0	+	+	0	0
	Valve Spool Grooving (F6.5)	0	0	-	-	0	-
Cylinder	Gas Enclosure (F7.1)	+/-	+/-	+/-	+/-	+/-	+/-
	Internal Leakage (F7.2)	+/-	+/-	+/-	+/-	+/-	+/-
Sensors	Flowmeter $\dot{V}$ (F20x)	0	0	0	0	+	0
	Pressure Chamber A p_A (F30x)	+	+	0	0	0	+
	Pressure Chamber B p_B (F40x)	0	0	-	-	0	-
	Pressure Valve Port P p_P (F501)	-	0	+	0	0	0
	Pressure at Pump p_S (F701)	0	-	0	+	0	0
	Flow Rate at Pump $\dot{V}_S$ (F80x)	0	0	0	0	0	0
	Piston Displacement y (F90x)	+	+	+	+	+	+
	Valve Spool Displ. y_V (F100x)	-	-	-	-	-	-

Faults Servo-Axis and Sensors

Fig. 5.10. Model-based fault-detection and diagnosis system (flow sensor has been investigated, but is not employed for fault detection)

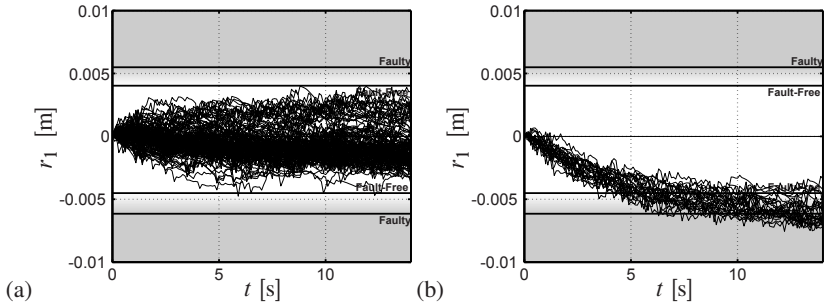


Fig. 5.11. Residual 1 for (a) fault-free case and (b) control edge erosion for several experiments

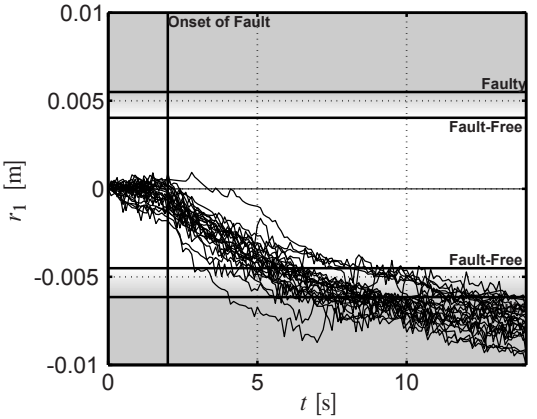


Fig. 5.12. Residual 1 for sensor offset in valve spool displacement $\Delta y_V = 0.002\,\mathrm{m}$

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{b}_{10}(T_\mathrm{P}) \\ \hat{b}_{11}(T_\mathrm{P}) \\ \vdots \\ \hat{b}_{14}(T_\mathrm{P}) \\ \hat{b}_{20}(T_\mathrm{P}) \\ \hat{b}_{21}(T_\mathrm{P}) \\ \vdots \\ \hat{b}_{24}(T_\mathrm{P}) \\ \hat{G}_{\mathrm{AB}}(T_\mathrm{P}) \\ \frac{1}{\hat{\tilde{E}}(T_\mathrm{P})} \end{pmatrix} \tag{5.1.21}$$

This parameter vector contains twelve parameters which may change with fluid temperature T_P. Although such a parameter estimation problem can easily be solved on modern computers, it will only be used for offline identification.

For online identification the parameter estimation problem is split into two parts since only specific combinations of flow paths can be active at any given time. Two independent parameter estimation problems for $y_V(k) \geq 0$ and $y_V(k) < 0$ are formulated to constrain the number of parameters to be estimated in each iteration. Since the valve spool displacement can only be positive or negative at any instant in time, the parameter estimation problem can be simplified.

For $y_V(k) \geq 0$, (5.1.9) can be written as

$$\begin{aligned} & b_{10}(T_P)\sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) \\ & + b_{11}(T_P)y_V(k)\sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) \\ & + \ldots + b_{14}(T_P)y_V(k)^4\sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) \\ & - G_{AB}(T_P)\,(p_A(k) - p_B(k)) - \frac{(V_{0A} + A_A y(k))}{\bar{E}(T_P)}\,\dot{p}_A(k) = A_A\dot{y}(k) \end{aligned} \tag{5.1.22}$$

Then a parameter estimation problem can be set up with the data matrix

$$\boldsymbol{\Psi}^T = \begin{pmatrix} \sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) & \ldots \\ y_V(k)\sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) & \ldots \\ \vdots & \\ y_V(k)^4\sqrt{|p_P(k) - p_A(k)|}\text{sign}\,(p_P(k) - p_A(k)) & \ldots \\ (p_A(k) - p_B(k)) & \ldots \\ A_A\; y(k)\;\dot{p}_A(k) & \ldots \end{pmatrix} \tag{5.1.23}$$

and the output vector

$$\mathbf{y}^T = (A_A\dot{y}(0), A_A\dot{y}(1), \ldots, A_A\dot{y}(k), \ldots, A_A\dot{y}((N-1))) \tag{5.1.24}$$

The solution of

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{\Psi}^T\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}^T\,\mathbf{y} \tag{5.1.25}$$

supplies estimates for the parameter vector $\vec{\theta}$ given as

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{b}_{10}(T_P) \\ \hat{b}_{11}(T_P) \\ \vdots \\ \hat{b}_{14}(T_P) \\ \hat{G}^+_{AB}(T_P) \\ \frac{1}{\hat{\bar{E}}^+_A(T_P)} \end{pmatrix} \tag{5.1.26}$$

where the superscript + denotes the estimates of the leakage coefficient that has been derived for a positive spool displacement. For $y_V(k) < 0$, a second, similar parameter estimation problem can be programmed. Then a parameter estimation problem can be set up with the data matrix

$$\boldsymbol{\Psi}^T = \begin{pmatrix} \sqrt{p_\mathrm{A}(k)} & \dots \\ y_\mathrm{V}(k)\sqrt{p_\mathrm{A}(k)} & \dots \\ \vdots & \\ y_\mathrm{V}(k)^4\sqrt{p_\mathrm{A}(k)} & \dots \\ (p_\mathrm{A}(k) - p_\mathrm{B}(k)) & \dots \\ A_\mathrm{A}\ y(k)\ \dot{p}_\mathrm{A}(k) & \dots \end{pmatrix} \tag{5.1.27}$$

The output vector

$$\mathbf{y}^T = (A_\mathrm{A}\dot{y}(0), A_\mathrm{A}\dot{y}(1), \dots, A_\mathrm{A}\dot{y}(k), \dots, A_\mathrm{A}\dot{y}((N-1))) \tag{5.1.28}$$

This second parameter estimation problem yields estimates $\vec{\theta}$ for

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{b}_{20}(T_\mathrm{P}) \\ \hat{b}_{21}(T_\mathrm{P}) \\ \vdots \\ \hat{b}_4(T_\mathrm{P}) \\ \hat{G}^-_{\mathrm{AB}}(T_\mathrm{P}) \\ \frac{1}{\hat{E}^-_\mathrm{A}(T_\mathrm{P})} \end{pmatrix} \tag{5.1.29}$$

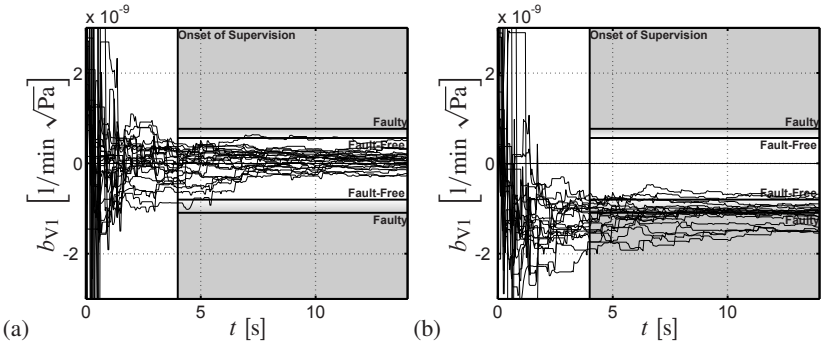


Fig. 5.13. Parameter estimation for a) fault-free case and b) control edge erosion

Figure 5.13 shows results of the parameter estimation for the fault-free case and for the case of control edge erosion. It can be seen from the diagrams that the parameter estimation reacts very sensitively to faults being present in the system.

As a way to compare the performance of different proposed fault-detection and diagnosis algorithms, the smallest reliably detectable sensor fault and the smallest reliably diagnosable sensor fault can be determined. For this purpose, the faults can easily be injected by post-processing of the measurements with varied fault size. These quantities have been determined for all sensors and are listed in Table 5.4.

Table 5.4. Smallest reliable detectable and diagnosable sensor faults

Sensor fault size	p_A	p_B	p_P	y	y_V
Minimum detectable	1.2 bar	0.4 bar	1.2 bar	3 mm	0.2%
Minimum diagnosable	3.7 bar	1.5 bar	3.0 bar	9 mm	0.4%

5.1.5 Conclusions

A precise modeling of the flow and cylinder is crucial for obtaining a sensitive model-based fault detection and diagnosis system. In order to increase the model fidelity, the differential equation is solved for the piston displacement instead of the chamber pressure. Furthermore, the valve characteristics should be approximated by a polynomial, which allows one to supervise proportional valves with an almost arbitrary geometry of the control edges. Finally, the temperature influence has to be taken into account in modeling, since even a small change of $\Delta T = 10°\,\mathrm{C}$ causes changes in the plant behavior precluding the detection of small faults. The use of parity equations for fault detection and diagnosis enables one to implement these algorithms on microcontrollers currently employed in hydraulic controls. The derived model can also serve as a fault-tolerant "model sensor" in the case of a loss of the piston displacement sensor.

By the aid of parameter-estimation methods, the fault diagnosis can be deepened. The parameter-estimation-based fault detection and diagnosis is very sensitive to process faults, but less for sensor faults if not specifically modeled for parameter estimation. Changes in the valve flow can be attributed precisely to erosion in one or more of the four control edges and valve spool grooving. By the estimation of the bulk moduli of the two chambers and the coefficient of internal leakage flow, it is in theory possible to detect gas enclosures (lowering the bulk modulus) and increased internal leakage (increasing the coefficient of internal leakage flow). Since the bulk modulus is estimated separately for chamber A and B, it is possible to pinpoint gas enclosures down to one of the two cylinder chambers. See Table 5.5 for a list of detectable and diagnosable faults depending on the sensor configuration.

These detailed investigations have shown that the detection of faults can be obtained with parity equations in realtime. In some cases this allows also fault diagnosis. After detection of the faults, a parameter estimation can be started in order to obtain additional symptoms which then allow one to diagnose most of the considered fault types.

5.2 Pneumatic actuators

Pneumatic actuators exploit the physical characteristics of compressed air. The high compressibility of air along with the capability of storing a larger amount of energy and the low viscosity of the transmission medium permit the design of efficient and fast drives. Offering a rugged and simple design (only one supply line necessary), these pneumatic drives are well suited for applications where typical forces of a few

Table 5.5. Detectable and diagnosable faults in dependence on the sensors used. Faults in parentheses are difficult to identify yet not completely impossible (depends on the individual setup)

Fault detectable	Sensor usage				
	y_V, y	p_S, y_V, y	$p_S, \dot{V}_S, y_V, y$	p_S, p_A, p_B, y_V, y	$p_S, \dot{V}_S, p_A, p_B, y_V, y$
Load must be constant	x	x	x		
Total breakdown	x	x	x	x	x
Pressure discharge valve			x		x
Supply-line congestion	x	x	x	x	x
Return-line congestion	x	x	x	x	x
Eroded control edge		(x)	(x)	x	x
Valve spool grooving		(x)	(x)	x	x
Gas enclosure		(x)	(x)	x	x
Internal leakage		(x)	(x)	x	x
Sensor faults	x	x	x	x	x

Fault diagnosable	Sensor usage				
	y_V, y	p_S, y_V, y	$p_S, \dot{V}_S, y_V, y$	p_S, p_A, p_B, y_V, y	$p_S, \dot{V}_S, p_A, p_B, y_V, y$
Load must be constant	x	x	x		
Total breakdown	x	x	x	x	x
Pressure discharge valve			x		x
Supply-line congestion		(x)	(x)	x	x
Return-line congestion		(x)	(x)	x	x
Eroded control edge		(x)	(x)	x	x
Valve spool grooving		(x)	(x)	x	x
Gas enclosure		(x)	(x)	x	x
Internal leakage		(x)	(x)	x	x
Sensor faults	x	x	x	x	x

N to some kN must be supplied. They can move at relatively high velocities and over long ranges. Besides these features, they are characterized by very safe operation even under extreme ambient conditions (temperature resistance, contamination resistance, overload capability, explosion-proof construction and fail-safe position through spring return). The pneumatic system is immune to interference caused by electric and magnetic fields, as well as radiation.

Pneumatic actuating devices can grossly be divided into pneumatic cylinders or diaphragms generating a translatory motion, and air motors generating a rotary motion. For more details see [5.2] and [5.8].

5.2.1 Pneumatic-actuator construction

Pneumatic actuators basically consist of an airflow valve and an actuating device, which transforms the pneumatic energy into mechanical energy. Figure 5.14 shows a scheme with an electrically controlled valve. The valve is connected to the pneumatic

pressure line which is supplied by an air compressor and manipulates the pressurized air flow to the actuating device.

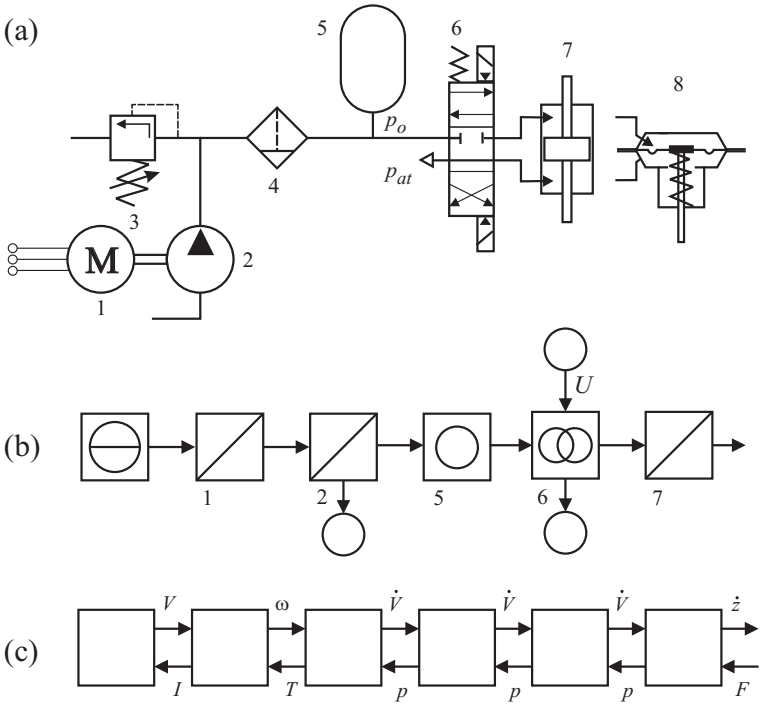


Fig. 5.14. Pneumatic actuator for linear motion with power supply: a) schematic; b) energy flow scheme; c) two-port representation. 1: AC motor, 2: air compressor, 3: pressure-relief valve, 4: air filter with water trap, 5: air storage (accumulator), 6: 4/3 proportional valve, electrical solenoid actuation, spring return, 7: double-rod cylinder; 8: diaphragm drive with spring return

The valves are either proportional-acting valves or switching valves. *Proportional-acting valves* allow a continuous manipulation of the airflow. They are manipulated either electrically by two solenoids as shown in Figure 5.14, or pneumatically with one air flow modulator and return spring, as depicted in Figure 5.15. *Switching valves* are electromagnetic devices and operate with pulse-width modulation (PWM) at high frequencies in order to generate a certain actuator position through the low-pass behavior of the cylinder diaphragm.

Pneumatic flow valves have in principle a design as shown in Figure 5.15. The pneumatic subsystem comprises a chamber sealed by a diaphragm that is acting on the valve stem. At the tip of the valve stem, a body (plug) is mounted which in conjunction with its counterpart, the valve seat, controls the hydraulic flow. Depending on the precision accuracy and the kind of fluid, different geometries are used. For very precise control tasks of small flows and large flows, usually needle-shaped bodies are used, whereas disk- or ball-shaped bodies are commonly chosen for fully

opening and closing valves. The stem passes through a gland (e.g., stuffing box) in order to seal the hydraulic system. Figure 5.15 shows a position controller (positioner), frequently mounted directly on the valve.

It varies the displacement of the valve stem in accordance with an external reference value. By means of a nozzle-flapper arrangement, the control error between the valve stem position and its reference value is sensed and the pressure supplied to the diaphragm chamber varied accordingly. Since the nozzle-flapper arrangement would not be able to supply a sufficient air flow, an air amplifier is connected in between the nozzle-flapper arrangement and the working chamber, allowing delivering or venting air flow.

These pneumatic flow control valves are produced in several different constructions. The valve bodies are, e.g. globe valves for straight pipelines with one or two plug stem guides, three-way valves or angle valves. The plug stem is usually sealed by a stuffing box (packing) or by metal bellows. The packing may consist of spring-loaded V-rings made of PTFE (polytetrafluorethylen), carbon, silk or graphite, depending on the liquid, steam, pressures and temperatures. The plug and seat are made of stainless steel and with or without stellite hard facing. Their design determines the k_V coefficient and the valve characteristic as well as the seat leakage. The normalized flow coefficient k_V according to DIN EN 60534 is defined for the flow of water as

$$k_V = \sqrt{\frac{\rho}{\rho_0}\frac{\Delta p_0}{\Delta p}}\,\dot{V} \tag{5.2.1}$$

with the standardized values $\rho_0 = 1000\,\mathrm{kg/m^3}$, $\Delta p_0 = 0.98\,\mathrm{bar}$ and water with dynamic viscosity of $\nu = 1\,\mathrm{mPa\,s}$.

k_{VS} corresponds to the rated full travel of $z = 100\%$. The valve characteristic determines the dependence between the normalized flow k_V in dependence on the travel z. It can be ordered with a linear or equal percentage (nonlinear) characteristic.

Figure 5.16 depicts a signal flow scheme for a pneumatic control valve with the position controller G_{C1} and the medium flow controller G_{C2}. The *position controller* is, as shown in Figure 5.15, traditionally a pneumatic proportional-acting controller, which manipulates the diaphragm chamber pressure p_1. Then the position z of the valve is measured mechanically and acts via a lever system on the amplifier. As it is not available as electrical signal it cannot be used for fault detection. If the *flow controller* G_{C2} is also pneumatic, the only electrical signals may be the output $\dot{V}$ of the flow rate transmitter and the setpoint W_2.

If, however, an electronic flow controller is used, the following signals may be available $\dot{V}$, W_2, W_1, U_1 and z.

5.2.2 Faults of pneumatic valves

Faults of pneumatic valves may arise in the air supply, air supply line, piston or diaphragm actuators, valve stem sealing and plug–seat combination. Table 5.6 lists some typical faults.

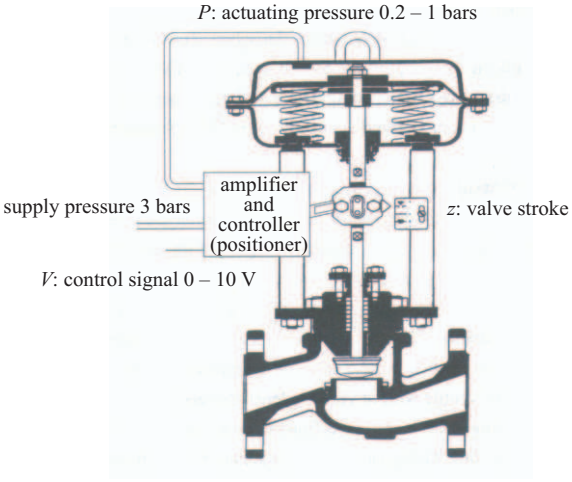

P: actuating pressure 0.2 – 1 bars
supply pressure 3 bars
amplifier and controller (positioner)
z: valve stroke
V: control signal 0 – 10 V

Fig. 5.15. Cross-sectional view of a pneumatic valve with an electro-pneumatic position controller

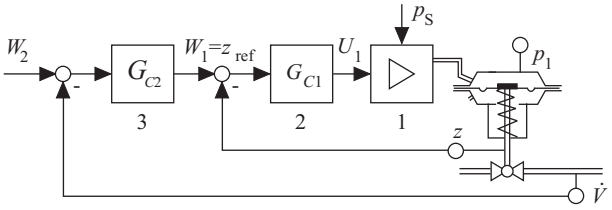


Fig. 5.16. Pneumatic flow valve with position and flow controller. p_1: diaphragm chamber pressure, p_s: supply air pressure; z: stem position; $\dot{V}$ flow rate; 1: amplifier (e.g., nozzle-flapper), 2: position controller; 3: flow rate controller

The possibilities of early fault detection depend strongly on the available electrical signals. If a pneumatic position controller is attached, the only electrical signals may come from a superimposed electronic controller, like a flow valve controller in the form of the flow rate $\dot{V}$, its reference value W_2 and the reference value W_1 for the positioner, compare Figure 5.16. Because the pneumatic position controller compensates some faults, only a few faults are detectable, see Table 5.6. If an additional electrical signal is added, like the chamber pressure p_1 the fault coverage improves. The situation improves further if the position controller is electronic, because then the manipulated variable U_1 and the valve position z are available, Table 5.6.

5.2.3 Models of pneumatic valves

To derive mathematical models of pneumatic valves, the same equation as for pneumatic cylinders, [5.9], can be used, if A_D is the area of the diaphragm and z the position of the valve stem. However, in contrast to pneumatic cylinders there exists

Table 5.6. Typical faults of pneumatic flow valves and fault-detection coverage in dependence on instrumentation. ×: yes, 0: no

		electronic flow controller		
		pneumatic positioner	electronic positioner	
valve		available measurements		
part	faults	$W_1, W_2, \dot{V}$	p_1	$W_1, U_1, W_2, z, \dot{V}$
pneumatic part	F_1 reduced supply pressure	0	×	×
	F_2 leakage in line to chamber	0	×	×
mechanical part	F_3 increased friction of plug stem	0	×	×
	F_4 increased valve cross-section area	×	×	×
	F_5 offset position sensor	×	×	×
flow sensor	F_6 offset flow sensor	×	×	×

only one chamber if the other is open to the atmosphere. The dry and viscous friction of the stuffing box and the plug stem guidance have a significant effect, see also [5.6], [5.17], [5.22]. [5.13] and [5.23] have shown how the (changing) friction can be identified and compensated by a model-based adaptive control algorithm.

The mass balance of the *pneumatic part*, the air chamber is

$$\begin{aligned} \dot{m}_1(t) &= \frac{d}{dt}\left(V_1(t)\rho_1(t)\right) = \dot{V}_1(t)\rho_1(t) + V_1(t)\dot{\rho}_1(t) \\ &= A_D\rho_1(t)\dot{z}(t) + \left(V_0 + A_D z(t)\right)\dot{\rho}_1(t) \end{aligned} \tag{5.2.2}$$

where V_0 is the vented chamber volume ($p_1 = 0$).

Using the gas equation $\rho = p/RT$ yields

$$\dot{p}_1(t) + \frac{A_D}{V_0 + A_D z(t)}\dot{z}(t)p_1(t) = \frac{RT_1}{V_0 + A_D z(t)}\dot{m}_1(t) \tag{5.2.3}$$

Hence, the parameters of this first-order differential equation for the chamber pressure are time-variant and depend on the motion and the position of the diaphragm.

The dynamics of the *mechanical part* of the valve, the diaphragm, the valve stem and the plug are governed by the balance of forces:

$$m_v\ddot{z}(t) + d_v\dot{z}(t) + c_s z(t) + f_c \operatorname{sign}\dot{z}(t) = A_D p_1(t) - F_{ext}(z) \tag{5.2.4}$$

where m_v is the mass of the stem and connected parts, c_s the spring constant, d_v and f_c the coefficients of the viscous and dry friction of the stuffing box and guidance and F_{ext} the external forces, mainly the force through the flowing fluid, which is proportional to the pressure drop Δp across the valve plug. The equation of the mechanical part is nonlinear because of the mostly considerable dry friction and the position-dependent plug force.

Both (5.2.3) and (5.2.4) lead to the signal flow diagram in Figure 5.17. The valve characteristic for the position follows from (5.2.4) with $d/dt \approx 0$:

$$z = \frac{1}{c_S}\left(A_D\, p_1 - F_{ext}(z) - f_c \operatorname{sign} \dot{z}\right) \tag{5.2.5}$$

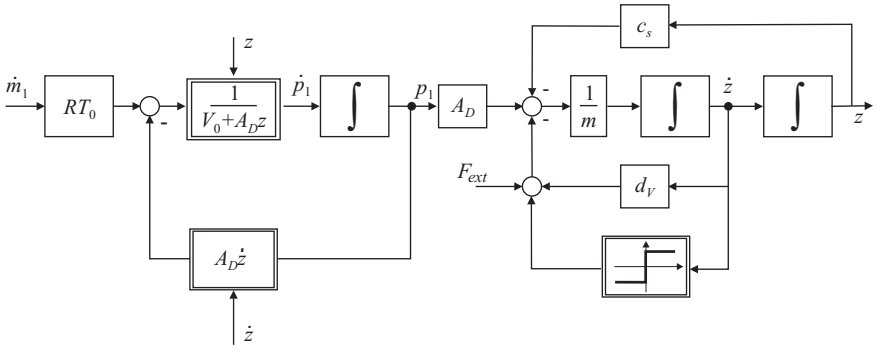


Fig. 5.17. Signal flow diagram of a pneumatic flow valve. $\dot{m}_1$: air mass flow into the diaphragm chamber

Neglecting the plug forces, the position z is linearly dependent on the chamber pressure, but with a direction-dependent dry-friction term, which results in a hysteresis curve. However, as measurements also show, this characteristic can be approximated by

$$z = z_0 - C_{1p}\, p_1 + f_{C-} = z_{00-} - C_{1p}\, p_1 \quad \text{for} \quad \dot{z} < 0 \tag{5.2.6}$$

$$z = z_0 - C_{1p}\, p_1 - f_{C+} = z_{00+} - C_{1p}\, p_1 \quad \text{for} \quad \dot{z} > 0 \tag{5.2.7}$$

with

$$z_{00-} = z_0 + f_{C-}; \qquad z_{00+} = z_0 - f_{C+} \tag{5.2.8}$$

This simplified valve characteristic holds at least piecewise over the whole travel way, as was shown by [5.6].

If the chamber pressure p_1 is not available as an electrical signal, then the manipulated variable U_1 of the position controller can be used as input of the valve. For a digital position controller this is for example the current output of an i/p-converter. The valve characteristics is then described by

$$z = z_{00-} - C_{1U} U_1 \quad \text{for} \quad \dot{z} < 0 \tag{5.2.9}$$

$$z = z_{00+} - C_{1U} U_1 \quad \text{for} \quad \dot{z} > 0 \tag{5.2.10}$$

Instead of U_1 also the reference value z_{ref} of the position controller can be taken.

Some references on the fault detection of pneumatic actuators are [5.11], [5.15], [5.16], [5.19], [5.24], [5.25]. The detection of stiction in treated in [5.5].

The following sections describe the results of some theoretical and experimental studies on the fault diagnosis of pneumatic valves for the case of pneumatic and electronic position controllers and different sensor usage and are based on [5.3], [5.6] and [5.7].

5.2.4 Fault detection with valve characteristics

The following experimental investigations are performed for a pneumatic valve with a pneumatic position controller and superimposed digital PI flowrate controller, see Figures 5.18 and 5.19. The flow rate control system is part of an industrial-scale thermal plant. First, the valve characteristics are considered for closed position loop and open flow control loop. Figure 5.20 depicts some characteristics for increased friction due to a too strongly tightened stuffing box. The characteristic $z(z_{ref})$ does not indicate significant differences because the position controller compensates the influence of higher friction to a certain degree. However, $p_1(z_{ref})$ and $z(p_1)$ show a significantly increased hysteresis behavior. The estimated dry friction is 2 to 3 times larger than the viscous friction and depends on the position z, [5.6].

Table 5.7 gives the influences of the considered faults on several characteristic curves, where approximate linear relations are assumed, as (5.2.5) to (5.2.9), and, e.g. Δz_{00} are offset changes and ΔC_1 are gain changes.

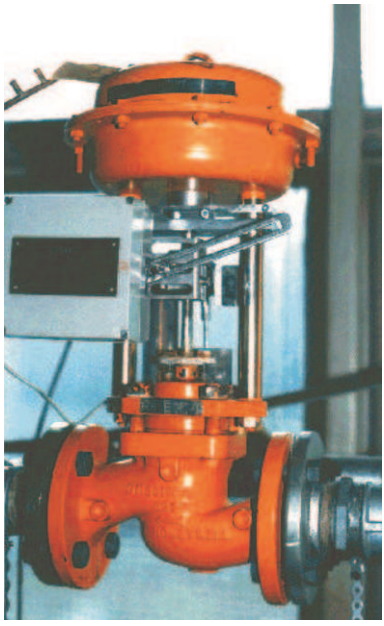


Fig. 5.18. Investigated electro-pneumatic valve (Honeywell type 2000, $K_{VS} = 25\,\mathrm{m}^3/\mathrm{h}$)

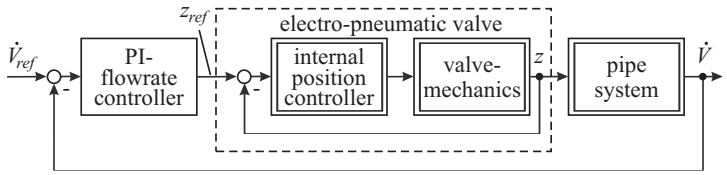


Fig. 5.19. Signal flow of the flow control loop with pneumatic valve and pneumatic position

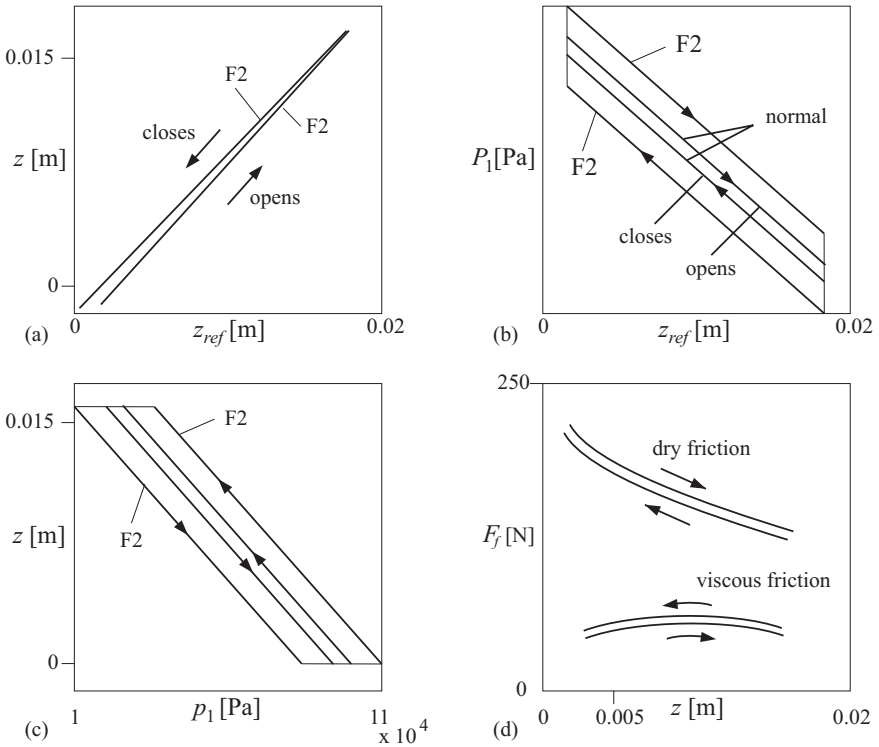


Fig. 5.20. Measured valve characteristic for normal and increased friction of the stuffing box (fault F2), [5.6]: a) $z = z_{ref}$, closed position loop; b) $p_1(z_{ref})$; c) $z(p_1)$; d) friction forces

Table 5.7. Fault-symptom table of valve characteristics for closed position and flow control loop

	$z(p_1)$		$z(z_{ref})$		$p_1(z_{ref})$		$\dot{V}(U_1)$	
faults	Δz_{00}	ΔC_{1p}	z_{00}	ΔC_{1u}	Δp_{10}	ΔC_{1Up}	$\Delta \dot{V}_{00}$	ΔC_{1V}
F1 supply-line leak	0	0	0	0	−	0	−	−
F2 increased friction	+ or −	0	0	0	+ or −	0	+ or −	0
F3 valve plug erosion	0	0	0	0	0	0	+	0
F4 position sensor error (+)	+	0	−	0	0	0	−	0
F5 increased pipe resistance	0	0	0	0	0	0	−	−
F6 flow sensor offset (+)	0	0	0	0	0	0	+	0

The different valve faults result in parallel shifts and/or slope (gain) changes of the characteristics. While $z(z_{ref})$ does not allow one to detect small faults the chamber pressure $z(p_1)$ and $p_1(z_{ref})$ are better suited to detecting some faults. The flow rate characteristic $\dot{V}(U_1)$ is influenced by all considered faults. However, the observed deviations depend also on changes in the connected plant. Hence, use of the chamber pressure p_1 and the flow rate $\dot{V}$ improve the possibilities for fault detection considerably.

For the identification of the characteristic curves either direction-dependent polynomials like

$$Y(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 \tag{5.2.11}$$

as for the hydraulic spool valve, Section 5.1, are recommended or local linear models, based on the LOLIMOT approach. Steady-state measurement can be gathered online during normal operations, if the plant operates with different flow rates. Alternatively, the valve characteristics can be determined by special test runs.

5.2.5 Fault detection of flow valves with pneumatic position controller

a) Measurement of fluid flow rate and reference value of position controller

If pneumatic valves have an electro-pneumatic transmitter to generate the pneumatic reference signals for the position controller and *no electrical valve position signal* z is available the behavior of the valve can only be observed through the resulting electrical signal from a flow or pressure measurement downstream in the plant. This case will be considered in the sequel first, because it is typical for chemical plants.

The task is to detect and diagnose faults in the pneumatic flow valve through measurements of the reference signal $W_1 = z_{ref}$ of the closed pneumatic position control loop, the controlled variable $\dot{V}$ and its reference value $W_2 = \dot{V}_{ref}$ of the superimposed flow control loop. Hence, methods for fault detection in closed loops are required. As shown in [5.10] different fault-detection methods have to be properly combined, e.g. features of closed-loop behavior, parameter estimation and parity equations. Parity equations for the controlled process can for linear closed loops be applied as in open loop.

For fault detection it is feasible to obtain the valve characteristic $\dot{V} = f(z)$ or $\dot{V} = f(p_1)$ which are nonlinear. As, however, z and p_1 are not measurable in this case the input–output behavior between the reference value z_{ref} and the resulting volume flow rate $\dot{V}$

$$\dot{V} = f(z_{ref}) \tag{5.2.12}$$

has to be used, which includes the behavior of the position control loop, see Figure 5.18, [5.4], [5.7]. Also this relation shows nonlinear behavior with dynamic and static characteristics depending on the operating point. Furthermore, the dynamics depend on the direction of the valve position changes (opening or closing).

This can be seen in Figure 5.21a) where the step response at distinct operating points is displayed. Note the different gain and time constants and the dead time of approximately 1 s. This open-loop behavior of the plant clearly indicates the strong

nonlinearities hindering the use of a normal linear model. The behavior of the valve operated in closed-loop control is shown in Figure 5.21b). The PI-controlled process shows fast behavior for valve opening (with spring force) and is significantly slower during valve closing (with pneumatic pressure force). Investigations yielded that a simplified linear first-order approximation of (5.2.12) for the closed-loop behavior around a constant local setpoint is sufficient. Therefore, local linear models can be used with operating-point- and direction-dependent parameters.

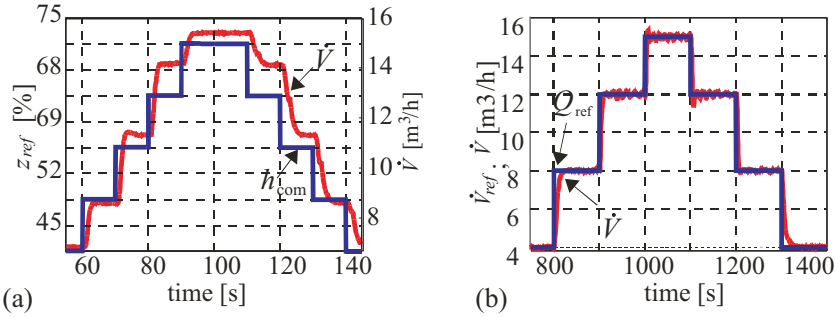


Fig. 5.21. Step responses of the pneumatic valve: a) open loop; b) closed loop

- *Fault detection and diagnosis with local linear models*
  As the valve shows nonlinear behavior it will be identified by using local linear models:

$$z(k) = w_0 + w_1 x_1(k) + \ldots w_n x_n(k) \tag{5.2.13}$$

  where $x_i(k)$ are different input signals and

$$w_j = \sum_{i=1}^{M} w_i \phi_i(\mathbf{x}) \tag{5.2.14}$$

  operating-point-dependent parameters, see [5.4], [5.9] and [5.28], with $\phi_i(\mathbf{x})$ membership functions depending on the input $\mathbf{x}$. This leads to a structure of the local linear process model shown in Figure 5.22.

  The symptoms are based on the comparison of features from the process with nominal ones from the model. For this purpose, a simulation with the model in closed-loop control (MCL) can be run in parallel with the process in closed-loop (PCL) using the same reference signal $W_1 = z_{ref}$. Figure 5.23 depicts this approach and the symptom generation. In this combination, three categories of symptoms can be distinguished, [5.7].

- *Parity equation residual*
  The residual-based symptoms can be derived in both closed-loop and open-loop operation, and normally do not require any additional process excitation. The

simplest residual is the output error between model and valve, within a time window of appropriate length l,

$$S_r = \frac{1}{l}\sum_{i=1}^{l} |\hat{z}_1(k-1) - z(k-1)|, \tag{5.2.15}$$

refer to Figure 5.23.

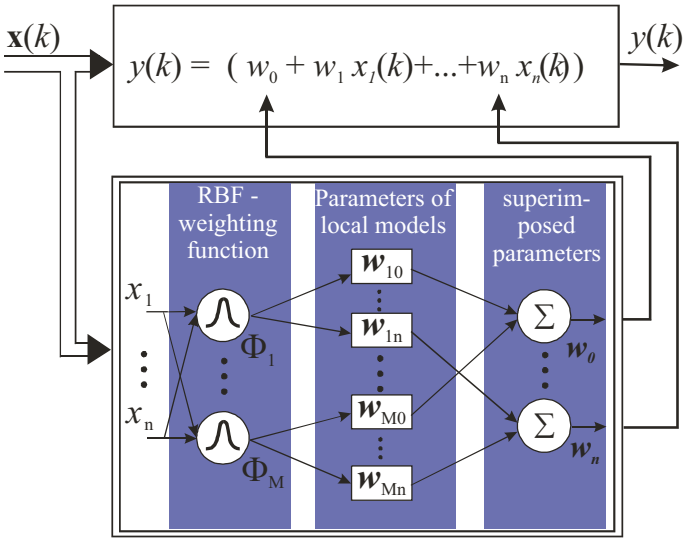


Fig. 5.22. Structure of the local linear model

- *Closed-loop performance-based symptoms*
  The performance of a control loop in constant operation regions and during set-point changes also characterizes the process since its performance will degrade should the system behavior change. Hence, symptoms can be derived by defining different control performance indices (CPI). One possibility is to rate the difference between the reference signal $W(k)$ and the controlled variable $y(k)$:

$$\begin{aligned} S_{CPI} &= I_{CPI} - \hat{I}_{CPI} \\ &= \frac{1}{l}\sum_{i=1}^{l}(W(k-1) - z(k-i))^2 \\ &\quad - \frac{1}{l}\sum_{i=1}^{l}(W(k-1) - \hat{z}_2(k-i))^2 \end{aligned} \tag{5.2.16}$$

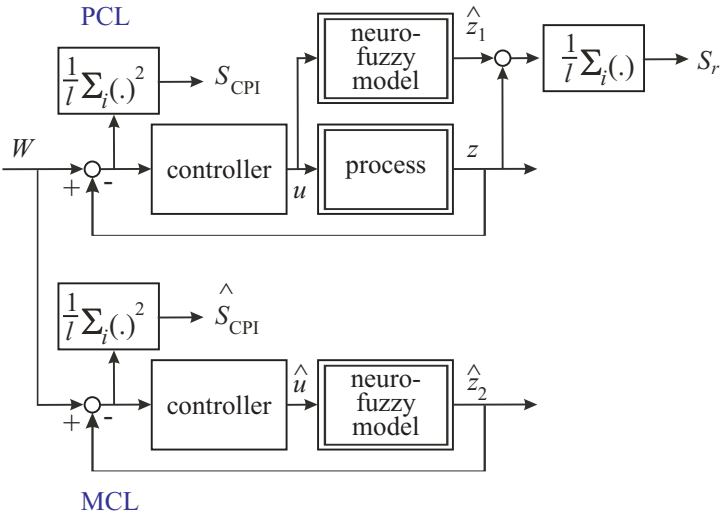


Fig. 5.23. Symptom generation using the process model

For constant reference values the symptom S_{CPI} depends not only on the control performance but also on disturbances and noise effects. Hence, to obtain significant symptoms, these noise effects must be relatively small.
To rate the control performance during setpoint changes S_{CPI} has to be divided by the setpoint difference, to make the symptoms comparable. Another possibility is the use of the time interval needed to enter a predefined region around the reference signal. In the following, this region is defined by 2% of the new setpoint. The symptom of the closed-loop behavior is named $S_{T98} = \hat{T}_{98model} - T_{98process}$.

- *Valve-parameter-based symptoms*
The third class of symptoms is derived from the parameters defining the process behavior. A common way is to identify a linear model in a small region around the actual setpoint. This requires sufficient process excitation which is often not fulfilled during normal operation. In order to meet this requirement, excitation signals can be added to the manipulated signal u, which leads to a temporary degraded control performance but enables a parameter estimation. The nominal values can be derived from the model of the fault-free process. Important features are the time constant T, gain K and offset O computed from the parameters of the discrete-time model of the valve. This yields the symptoms, [5.7]

$$\begin{aligned} S_T &= \frac{\hat{T}}{T_{dyn.lin.}} \\ S_K &= \frac{\hat{K}}{K_{dyn.lin.}} \\ S_O &= \hat{O} - O_{dyn.lin.} \end{aligned} \tag{5.2.17}$$

These parameters can also be extracted from local linear models of the closed-loop behavior with z_{ref} as input and z as output, [5.3].
The next step is the determination of fault symptom relationships. This can be solved by prior knowledge and physical considerations or from experiments.

- *Fault diagnosis*
  The evaluation of the fault symptom patterns is performed using a generated fuzzy classification tree, using the self-learning classification tree (SELECT) method, [5.7], [5.10]. Its advantages for the classification of fault symptoms are the transparency of the resulting classifier, the simple integration of a priori knowledge as fuzzy rules and the intuitive concept.

  The approach is divided into five phases:

  1) Building appropriate fuzzy membership functions (MSF) and removal of irrelevant symptoms
  2) Selecting a rule for the easiest separable fault
  3) Removal of measurements of this fault from the data set
  4) Back to Step 2 until all faults are considered, thereby creating a tree structure
  5) Fine tuning of relevance weights by constrained optimization.

- *Experimental results*
  As approximation of (5.2.12) a discrete-time neuro-fuzzy model with 12 local models has been identified. The sample time is $T_0 = 0.1\,\mathrm{s}$. Each local model is a first-order transfer function with a dead time of $T_d = 1\,\mathrm{s}$ or $d = T_d/T_0 = 10$. The LPV (linear parameter variable) model structure is

$$\dot{V}(k) = w_0 + w_1 z_{ref}(k-1-d) + w_2 \dot{V}(k-1) \tag{5.2.18}$$

  where the parameters depend on the reference signal and the direction of the valve motion:

$$w_i = f(z_{ref}(k-d),\ \mathrm{sign}\,(z_{ref}(k-d))) \tag{5.2.19}$$

  The task is to detect and diagnose faults in closed-loop operation of the flow circuit. The investigated faults are:

  F1: Leak in the pneumatic pipe between position controller and membrane chamber.
  F2: Increased friction of the valve (stuffing box).
  F3: Erosion of the valve plug (simulated by a bypass).
  F4: Fault in the valve position controller.
  F5: Increased flow resistance of the pipe system (partial clocking).
  F6: Fault in $\dot{V}$ sensor ($\dot{V}_{sensor} = 1.1\dot{V}_{normal}$).

  The generation of the symptom S_T, S_K, S_O requires sufficient process excitation for online valve estimation of the linear parameter model in the actual setpoint. In typical applications this is often not fulfilled, due to constant setpoints and the slow dynamics of the position control. In order to estimate the parameters from (5.2.18), an additional input excitation is required. In this approach, two small steps of 10 s duration are added to z_{ref} as can be seen in Figure 5.24. It

shows the reference small signal and the control performance degradation of the flow through the valve. Herewith, the average flow is kept constant. The faults effect the symptoms in different ways. The influence of selected faults on three symptoms is depicted in Figure 5.25.

Using all six symptoms, a classification using the SELECT method was performed. The total training data consisted of 60 measurements for each of the six faults. For learning and validation, a three-fold cross-validation scheme was employed. However, there exists some overlap of some of the fault classes as depicted in Figure 5.25. A classification rate of 100% cannot therefore be expected. Additionally, it was not possible to train only one diagnostic tree for the complete range of operation. Some of the symptoms show distinct differences in their behavior for small and large flow rates. Therefore, two trees were learned, one for high flow rates (above 50% of z_{ref}) and one for lower flow rates. Here, only results of the higher flow rates are presented. The results for flow rates below 50% are similar.

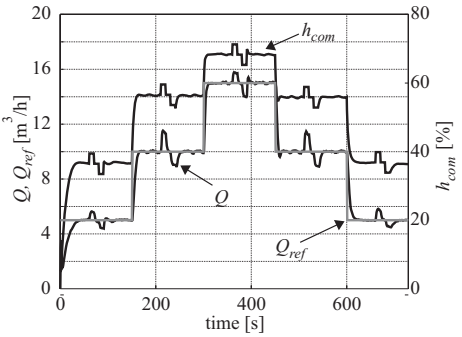


Fig. 5.24. Step responses of the flow control loop with additional test signals of the position reference value for process excitation

Table 5.8 displays the fuzzy relationships between the faults and symptoms as they were extracted from the classification tree. It can be seen that symptoms S_K and S_O are more included for the classification than symptom S_T for instance. This indicates their relevance for the distinction of the faults. Table 5.8 also ranks the faults according to the difficulty of separation. Clearly, the erosion and leak in the supply air pipe system seem to be difficult to classify, whereas flow resistance and controller faults can be found more easily. The achieved classification rate was 81% when all six faults were considered. A diagnosis without the erosion fault resulted in a classification rate of 84%. With less stringent requirements for the diagnosis even higher values are possible. Additionally, the diagnosis system also produces membership values for the faults which were not detected thereby giving an indicator for alternative possible fault causes. For more details see [5.4].

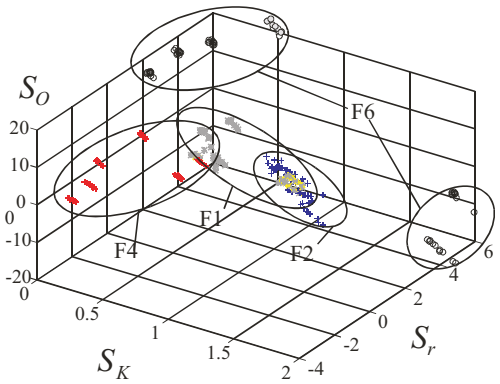


Fig. 5.25. Reduced symptom space for some faults

Table 5.8. Fuzzy symptoms at the nodes – AND relations of the classification tree

faults	S_K	S_r	S_O	S_T	S_{T98}	S_{CPI}
F5 increased pipe resistance	not large	large			changed	
F4 position sensor error			changed		unchanged	changed
F2 increased friction				large	unchanged	unchanged
F6 flow sensor offset	large		small			unchanged
F3 valve plug erosion	not small					unchanged
F1 supply leak						changed

Hence, by identifying local linear direction-dependent valve models and generation of symptoms from output parity equations, valve model parameters and position control performance it is possible to detect and diagnose at least six typical faults of pneumatic valves and its flow control. The required measured signals are the fluid flow rate $\dot{V}$, its reference value $\dot{V}_{ref}$ and the position reference value Z_{ref} of the valve. Electrical signals of the valve position z and the diaphragm pressure p_1 were not used. Therefore, this fault-diagnosis procedure is suitable for flow control with pneumatic valves and pneumatic position controllers, frequently used in chemical plants. However, the identified dynamic models for the flow rate depend on the connected plant and may be influenced by changes of its parameters or structure. The local linear models can be tuned during a setup phase at commissioning by a test run of the flow control.

b) Measurement of valve position and chamber pressure

For the same pneumatic flow valve as described in the last section now only the valve position z and its reference value Z_{ref} is used, [5.3]. Local linear models with z-transfer function

$$G_1(z) = \frac{z(z)}{z_{ref}(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-4} + b_3 z^{-7}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} + c_0 \qquad (5.2.20)$$

are identified with linear parameter variable models. Ten models were used with additional input by the first derivative of $z_{ref}(z)$, in order to take direction-dependent behavior into account. Based on this model for the position control loop and excitation by reference step inputs the following characteristic features were extracted: gain S_K, dominant time constant S_T and DC-parameter (offset) S_O as given in (5.2.17), and output residual S_r as (5.2.15). The resulting fault-symptom table is shown in Table 5.9. Because of the different patterns the investigated faults except F3 can be isolated and therefore diagnosed. However, the symptom generation depends on the operation point of the valve. For example, plug erosion (larger cross-sectional area between plug and seat) can be better detected for more closed valve position and increased pipe resistance for more open valve position.

Hence, the measurement of the valve position and the position controller reference value allows to diagnose several faults by nonlinear model-based fault detection with parameter estimation and external stepwise excitation of the position controller setpoint.

Table 5.9. Fault-symptom table of the valve position control loop with parameter estimation

	symptoms			
faults	closed-loop gain S_K	closed-loop time constant S_T	offset parameter S_O	output residual S_r
F1 supply leak	−	0	0	+
F2 increased friction (stuffing box)	0	+−	++	−−
F3 valve plug erosion	+	−	−−	+
F4 position sensor error	−−	−−	++	++
F5 increased pipe resistance	++	++	−−	−
F6 flow sensor offset	0	+−	0	0−

c) Additional measurement of chamber pressure and pressure drop

Now, a case is considered where no external input is required. If in addition to the measured signals z, z_{ref}, $\dot{V}$, $\dot{V}_{ref}$ also the diaphragm pressure p_1 and pressure drop Δp over the valve and available parity equations are established. They use difference equation models with z-transfer functions

$$G_1(z) = \frac{z(z)}{z_{ref}(z)}; \quad G_2(z) = \frac{\dot{V}(z)}{z(z)}; \quad G_3(z) = \frac{\dot{V}(z)}{z_{ref}(z)}$$
$$G_4(z) = \frac{p_1(z)}{z_{ref}(z)}; \quad G_5(z) = \frac{\Delta p(z)}{z(z)} \tag{5.2.21}$$

which are identified for normal status with six to ten local linear models of order $n = 1 \ldots 3$, [5.3]. Based on these fixed multi-models the following residuals of parity equations are calculated:

$$
\begin{aligned}
r_1(k) &= z(k) - G_1 \cdot z_{ref}(k) \\
r_2(k) &= \dot{V}_{ref}(k) - G_2 \cdot z(k) \\
r_3(k) &= \dot{V}_{ref}(k) - G_3 \cdot z_{ref}(k) \\
r_4(k) &= p_1(k) - G_4 \cdot z_{ref}(k) \\
r_5(k) &= \Delta p(k) - G_5 \cdot z(k)
\end{aligned}
\tag{5.2.22}
$$

The resulting deflection of these residuals for implemented faults during normal operation is stated in Table 5.10. Except F2 and F4 the remaining four faults can be isolated.

Hence, the additional measurements allow one to use more parity equations and do not require permanent parameter estimation and extra excitation, as in the last section. However, the dynamic models used should be adapted to eventually changed plant behavior after certain time intervals.

Table 5.10 shows that if only measurements of z, z_{ref} and p_1 are available two residuals r_1 and r_4 can be used. Then fault F5 and two groups of faults can be detected, F1 or F5 and F2 or F4.

Further methods and experimental results for the fault diagnosis of pneumatic valves were investigated by [5.6]. Through measuring z_{ref}, z, $\dot{V}$ and p_1, the actuating pressure in the diaphragm chamber, a combination of parameter estimation for linear models and closed-loop characteristics allowed to diagnose several faults.

Table 5.10. Fault-symptom table of the valve with parity equations and additional measurements. 1: residual deflected, 0: no change

	symptoms				
faults	r_1	r_2	r_3	r_4	r_5
F1 supply leak	1	0	0	0	0
F2 increased friction (stuffing box)	1	1	1	1	1
F3 valve plug erosion	0	1	1	0	0
F4 position sensor error	1	0	1	0	1
F5 increased pipe resistance	1	1	1	0	1
F6 flow sensor offset	1	1	1	1	1

5.2.6 Fault detection of flow valves with electronic position controller

The use of digital position controllers instead of pneumatic ones allows the access to electrical signals of the position reference variable $W_1 = z_{ref}$, the valve stem position z and a signal of the manipulated variable U, compare Figure 5.16. As indicated in Tables 5.6 and 5.7 this enables a good coverage for the detection of faults in the pneumatic and mechanical parts of the valve. As then an input and an output variable of the valve is available, model-based fault-detection methods can be applied, as shown for the hydraulic servo axis, Section 5.1 and in [5.6].

However, the manipulating pressure p_1 acting on the diaphragm after the air amplifier is usually not measured but only an electrical signal, generating the input pressure to the amplifier. This value is for example available from the i/p converter of the position controller where i is an applied standard DC current of $4 - 20\,\text{mA}$ and is a measure of the manipulated variable U_1 of the position controller.

In [5.15], [5.16] it was described how a digital position controller provides additional information about its own condition as well about the pneumatic valve. The positioner is still realized as an analog PD-controller with analog position signal and analog manipulated variable for an i/p converter, but obtains its reference value from a microprocessor. Standard diagnosis functions then give information on the device status (e.g. run time, configuration data, exceeded limits) and on the operation (zero point, hardware faults, data faults). An extended diagnosis allows the evaluation of the measurements. Histograms give statistics on the position operating area and diagrams between the valve position and the manipulation DC current i. This gives hints to leakages, changed spring forces, supply pressure and backward forces from the flow medium. Through the application of small changes of the reference variable W_1 and observation of the stem position, increased friction of the plug stem can be detected. However, a detailed fault diagnosis seems not to be possible. For an improvement further sensors (pressures, structure-borne noise) are proposed.

A digital position controller and a valve stem potentiometer [5.11] are used to train a multi-layer neural network that learns the relationships between seven characteristic figures of position step responses, like dead time, rise times, overshoot and steady-state error for three faults, incorrect supply pressure, vent blockage of one diaphragm chamber and diaphragm leakage. The valve is taken offline (no process flow) and step responses are only measured in one direction, avoiding effects of hysteresis and dead band. The three faults could be detected and isolated, but other faults like valve stem friction are not covered.

5.2.7 Conclusions

The different approaches for fault detection and diagnosis of pneumatic valves have shown that the fault detection and diagnosis coverage depends strongly on the instrumentation of the valve. It could be shown that for a *pneumatic position controller* with the position setpoint as the only available electrical signal, the superimposed (cascaded) electronic flow control gives access to the flow rate and its setpoint and therefore allows one to diagnose several valve and also flow control faults. If additional measurements are available, such as valve position or chamber pressure the identification of static characteristics or the use of nonlinear discrete-time (local linear) models for the closed loop allow one a more detailed fault detection of the valve and its control. In the last case parameter estimation and setpoint changes such as excitation or parity equations in normal operation can be applied.

If *digital position controllers* are implemented, then usually the valve stem position and the controller output can be used. Model-based fault-detection methods based on static characteristics, parity equations and parameter estimation then allow a specific valve fault diagnosis, in addition to the conventional approaches, like limit

checking of supply air pressure, histograms of valve position, observation of limit switches, etc.

Part III

Machines and Plants

6

Fault diagnosis of pumps

Pumps are basic components in most technical processes, like in power and chemical industries, mineral and mining, manufacturing, heating, air conditioning and cooling of engines. They are mostly driven by electrical motors or by combustion engines and consume a high percentage of electrical energy. One distinguishes mainly centrifugal pumps for high deliveries with lower pressures and hydrostatic or positive displacement (reciprocating) pumps for high pressures and small deliveries. They transport pure liquids, or mixtures of liquids and solids and herewith increase the pressure to compensate, e.g. for resistance losses or enabling thermodynamic cycles.

In the past, circular pumps were mostly driven with constant speed and the flow rate of liquids was manipulated by valves with corresponding throttling losses. Due to the availability of cheaper speed-controlled induction motors also circular pumps with lower power are now used for controlling the flow rate in order to save energy.

The overall reliability and safety of many plants depends on the health of pumps. Therefore, the supervision and fault diagnosis of pumps is of relatively high importance. In this chapter the results of several case studies will be treated for centrifugal pumps and reciprocating pumps.

6.1 Centrifugal pumps

6.1.1 State of the art in pump supervision and fault detection

Damage to centrifugal pumps occurs either in the hydraulic parts or the mechanical parts. An inquiry by the German Fachgemeinschaft Pumpen, VDMA [6.29], mainly among chemical industry and water treatment plants has shown the following result: 59% of pumps operate continuously, 19% daily and 22% for a short time. Inspection intervals are three months on average. Unplanned repairs because of defects happened within a mean of nine months.

Table 6.1 shows the faulty components as causes of damages to centrifugal pumps.

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_6, © Springer-Verlag Berlin Heidelberg 2011

The most frequent faulty components are therefore the sliding ring seals and the ball bearings, see Figure 6.1. Causes for faults which lead at least to interruptions of operation or maintenance are shown in Table 6.2. Cavitation, dry run, increased wear and deposits are especially important for fault detection.

Table 6.1. Faulty components as cause for damages to centrifugal pumps, [6.29]

Faulty components	Reported frequency [%]	Faulty component	Reported frequency [%]
sliding ring seal	31	sliding bearings	8
rolling bearing	22	clutch	4
leakage	10	split pipe	3
driving motor	10	casing	3
rotor	9		

Table 6.2. Malfunctions of centrifugal pumps and their explanation

Faults	Explanation and consequences
cavitation	development of vapor bubbles inside the fluid if static pressure falls below vapor pressure. Bubbles collapse abruptly leading to damage at the blade wheels and generate crackling sound
gas in fluid	A pressure drop leads to appearance of solved gas in the transported liquid. A separation of gas and liquid and lower head may result
dry run	missing liquid leads to lack of cooling and overheating of bearing. Important for starting phase
wear	*erosion*: mechanical damage to walls because of hard particles or cavitation *corrosion*: by aggressive fluids *bearings*: mechanical damage through fatigue and metal friction, generation of pittings and rents *plugging of relief bore holes*: leads to overloading of axial bearings and their damage *plugging of sliding ring seals*: leads to higher friction and smaller efficiency *increase of split seals*: leads to less efficiency
deposits	deposits of organic material or through chemical reactions at the rotor entrance or outlet lead to less efficiency, higher temperatures until total breakdown of pumping
oscillations	unbalance of the rotor through damage or deposits at the rotor, damage to the bearings

The supervision of pumps depends very much on the applied *instrumentation*. If for example only the outlet pressure is measured for a centrifugal pump with constant known speed only large deviations to the normal operation pressure give hints

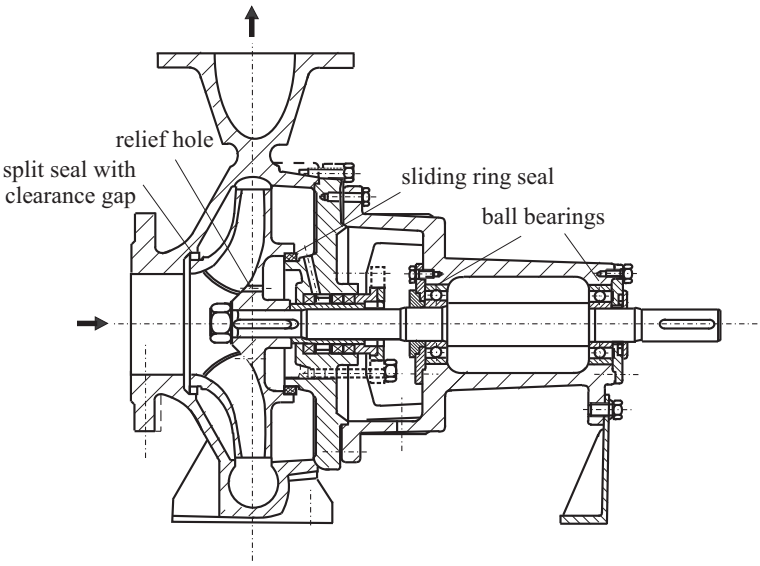


Fig. 6.1. Sectional drawing of a typical centrifugal pump (KSB Etanorm)

that somewhere large faults or failure happened. The addition of volume flow rate measurement improves the situation, as then changes of the head characteristic can be observed, however, without the possibility of fault diagnosis. The measurement of the inlet pressure allows one to observe the NPSH (net positive suction head) with regard to cavitation. All these simple supervision methods generally do not allow an early detection of small faults and do not give information on the causes of faults (diagnosis).

The supervision of *well-instrumented pumps* is usually based on the measurement of inlet and outlet pressure or the head only, the flow rate, speed and temperature of bearing casing and limit checking of these values. For example, if the outlet pressure or flow rate are too low (or too high) compared to the normal or rated values, this may be the result of gas enclosures, dry run, large deposits, strong deposits or bearing or motor defects. These large deviations from normal operation are easily observable by the exceeding of adjusted thresholds. But a fault diagnosis and an early fault detection is in general not possible with this checking of limits.

Various research efforts have given an insight into the pump behavior under the influence of faults. The application of *vibration sensors* and analysis of structure-borne noise is investigated, e.g. by [6.18], [6.10], [6.25], [6.20], [6.17]. The methods require special sensors and pump-specific and signal-specific evaluation methods and allow one to detect vibration-related faults under certain operating conditions, see the discussion in Section 6.1.5.

The simultaneous evaluation of several measurements and development of *model-based fault-detection methods* was performed by [6.6] (pump with variable speed and parameter estimation), [6.22] (pump with constant speed and parameter esti-

mation), [6.5] and [6.30] (pumps with different power and combination of parity equations and parameter estimation). These research projects are the basis for the following sections. Further publications on model-based approaches are [6.8] confirming parameter-estimation-based approaches, and [6.1] applying a state observer and parameter estimation.

6.1.2 Models of centrifugal pumps and pipe systems

a) Pump

The torque M applied to the rotor of a radial centrifugal pump leads to a rotational speed ω and transmits a momentum increase of the liquid from the rotor inlet with smaller radius r_1 to the rotor outlet with larger radius r_2 by guiding the liquid through blade-bounded channels. The theoretical required torque follows from an angular-momentum balance equation, known as Euler's turbine equation. This leads to the theoretical pump head

$$H_{th} = h_{th1}\omega^2 - h_{th2}\omega\dot{V} \tag{6.1.1}$$

where the delivery head is defined as

$$H = \frac{p_2 - p_1}{\rho g} = \frac{\Delta p}{\rho g} \tag{6.1.2}$$

with p_1 the pressure at the inlet and p_2 at the outlet. $\dot{V}$ is the volume flow rate.

Taking into account a finite number of blades, blade and tube friction losses, impact losses due to nontangential flow at the blade entrance, the basic equation for the delivery head of the pump becomes, [6.2], [6.5], [6.26]:

$$H = h_{nn}\omega^2 - h_{nv}\omega\dot{V} - h_{vv}\dot{V}^2 \tag{6.1.3}$$

The coefficients h_i are determined by the basic equations and contain empirically determined parameters.

The corresponding power transmitted to the fluid is

$$P = \rho g H\dot{V} = M\omega \tag{6.1.4}$$

The theoretical pump torque then results with (6.1.2) and (6.1.1):

$$M_{th} = \rho g\frac{\dot{V}}{\omega}H_{th} = \rho g\left(h_{th1}\omega\dot{V} - h_{th2}\dot{V}^2\right) \tag{6.1.5}$$

Including the flow losses, (6.1.3) has to be inserted in (6.1.4), resulting in the real torque:

$$M_P = \rho g\left(h_{nn}\omega\dot{V} - h_{nv}\dot{V}^2 - h_{vv}\frac{\dot{V}^3}{\omega}\right) \tag{6.1.6}$$

The mechanical part of the pump is modeled by the rotational impulse balance

$$J_P \frac{d\omega(t)}{dt} = M_{mot}(t) - M_P(t) - M_f(t) \tag{6.1.7}$$

where J_P is the ratio of inertia of motor and pump and M_f is the friction torque, consisting of Coulomb friction M_{f0} and viscous friction $M_{f1}\omega$:

$$M_f(t) = M_{f0} \operatorname{sign} \omega(t) + M_{f1}\omega(t) \tag{6.1.8}$$

Figure 6.2 shows the resulting characteristics of the delivery head H and the torque M of a centrifugal pump.

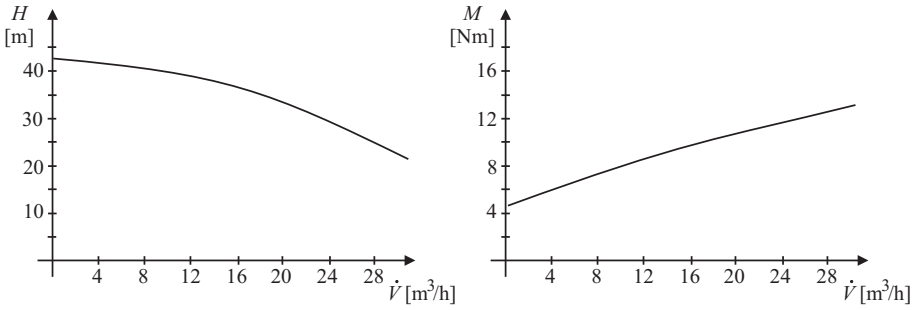


Fig. 6.2. Measured characteristics of a centrifugal pump ($n_N = 2900$ rpm)

b) Pipe system

The pump is now assumed to transport a fluid in a pipe system from a lower to an upper storage tank, Figure 6.3. Based on the momentum balance equation of the pipe one can state for turbulent flow

$$H(t) = a_F \frac{d\dot{V}(t)}{dt} + h_{rr}\dot{V}^2(t) + H_{stat} \tag{6.1.9}$$

where h_{rr} is a resistance coefficient of the pipe, taking into account pipe parts, pipe elbows and throttle valves, $a_F = l/gA$ with l the pipe length and A the pipe cross-sectional area, and H_{stat} is the height of the storage over the pump (resp. lower storage height), which means that the static pressure is

$$p_{stat} = \rho g \; H_{stat} \tag{6.1.10}$$

Assuming that (6.1.9) can be linearized for small deviations around the steady-state volume flow $\bar{\dot{V}}$ and $p_1 =$ const. leads to a first-order differential equation

$$T_F \frac{d\Delta\dot{V}(t)}{dt} + \Delta\dot{V}(t) = K_F \Delta p_2(t) \tag{6.1.11}$$

with gain and time constant

$$K_F = \frac{\Delta p_2}{\Delta \dot{V}} = \frac{1}{2\rho g h_{rr} \bar{\dot{V}}} \tag{6.1.12}$$

$$T_F = \frac{\alpha_F}{2h_{rr}\bar{\dot{V}}} = \frac{l}{2gAh_{rr}\bar{\dot{V}}} \tag{6.1.13}$$

Hence, the time constant of the fluid increases with larger pipe length l and smaller volume flow $\dot{V}$.

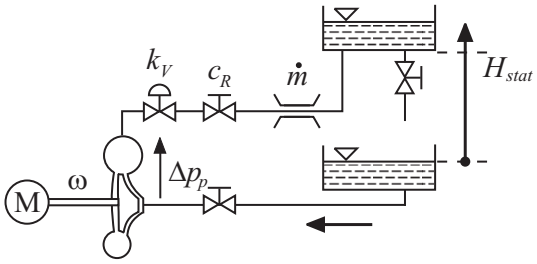


Fig. 6.3. Scheme of a circular pump with electrical drive, pipe system and storage. Open circuit

c) Pump and closed pipe circuit system

If the pump transports the fluid through a closed pipe circuit system with constant resistance parameters, as shown in Figure 6.4, the steady-state behavior follows from (6.1.3) and (6.1.9) with $d\dot{V}/dt = 0$ and $H_{stat} = 0$:

$$(h_{rr} + h_{vv})\,\dot{V}^2 + h_{nv}\omega\dot{V} - h_{nn}\omega^2 = 0 \tag{6.1.14}$$

The solution of this quadratic equation yields

$$\dot{V} = \kappa\omega \tag{6.1.15}$$

with

$$\kappa = \frac{1}{2(h_{rr} + h_{vv})}\left(-h_{nv} + (-)\sqrt{h_{nv}^2 + 4h_{nn}(h_{rr} + h_{vv})}\right) \tag{6.1.16}$$

If the resistance parameters in the pipe are constant (i.e. no valve position changes) the volume flow is proportional to the speed of the pump.

Because of the nonlinear pump torque the speed dynamics (6.1.7) are also nonlinear. However, for small deviations $\Delta\omega$ around a steady state $\bar{\omega}$ the equation can be linearized. In the case of a closed pipe circuit (6.1.15) holds and the pump torque (6.1.6) reduces to

$$\begin{aligned} M_P &= \rho g \kappa \left(h_{nn} - h_{nv}\kappa - h_{vv}\kappa^2 \right) \omega^2 \\ &= k_m \omega^2 \end{aligned} \tag{6.1.17}$$

Then (6.1.7) changes to

$$J_P \frac{d\omega(t)}{dt} = M_{mot}(t) - k_m \omega^2(t) - M_{f0} \operatorname{sign} \omega(t) - M_{f1}\omega(t) \tag{6.1.18}$$

Linearizing yields the first-order differential equation

$$T_P \frac{d\omega}{dt} + \Delta\omega(t) = K_P \Delta M_{mot}(t) \tag{6.1.19}$$

with the gain

$$K_P = \frac{\Delta\omega}{\Delta M_{mot}} = \frac{1}{2k_m\bar{\omega} + M_{f1}} \tag{6.1.20}$$

and the time constant

$$T_P = \frac{J_P}{2k_m\bar{\omega} + M_{f1}} \tag{6.1.21}$$

The time constant of the motor and pump gets small for small ratio of inertia J_P and large speed and has to be compared to the fluid time constant T_F due to (6.1.13). Figure 6.5 shows a signal flow diagram for the linearized pump–pipe system, taking into account (6.1.11) and (6.1.19). Hence, a second-order dynamic system results, according to the momentum stored by the pump-motor rotor and the fluid mass in the pipe.

6.1.3 Fault detection with parameter estimation

Model-based fault detection with parameter estimation is described first for the static behavior and then for the dynamic behavior of a centrifugal pump. Figure 6.4 shows the scheme of the investigated pump and the pipe circuit.

a) Constant speed operation and shut-off

In many cases an AC motor without speed control is used as a pump drive. [6.22], [6.23] investigated two possibilities. By changing the position of a valve in the piping system the plant characteristic changes and therefore the operating point of the pump, which is determined by the intersection between pump and plant characteristic. Thus, the parameters of the static head-flow curve $H(\dot{V}, \omega)$ can be identified. Except for the small changes due to the slip-torque characteristic of the AC motor, the pump speed remains constant. A second possibility consists in the evaluation of the dynamic pump behavior at shut-off. Thereby, the pump is shut off while operating at a normal flow or at zero flow. Both cases were examined. The following signals were measured: U, AC motor voltage; $\dot{V}$, volume flow; I, AC motor current, H, pump head; ω, speed.

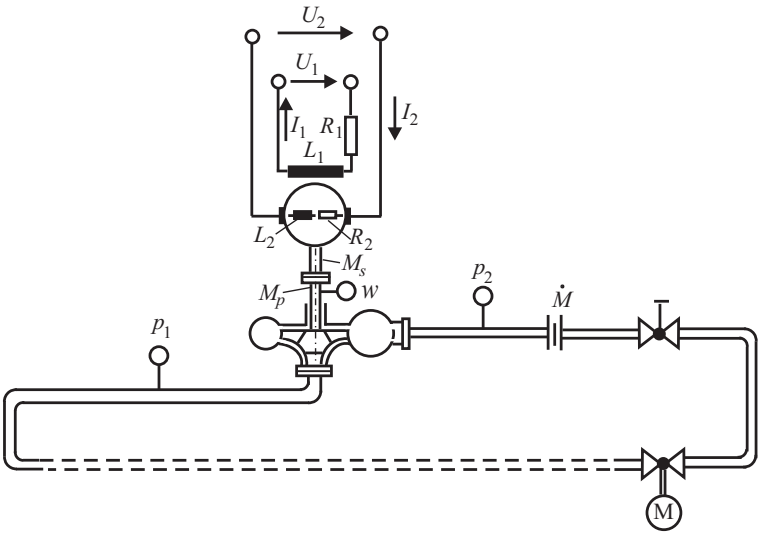


Fig. 6.4. Scheme of the speed-controlled DC motor and centrifugal pump. Closed circuit. Motor: $P_{max} = 4\,\text{kW}$; $n_{max} = 3000\,\text{rpm}$; pump: $H = 39\,\text{m}$; $\dot{V}_{max} = 160\,\text{m}^3/\text{h}$; $n_{max} = 2600\,\text{rpm}$. An AC motor was used for steady-state operation and a DC motor for dynamic operation

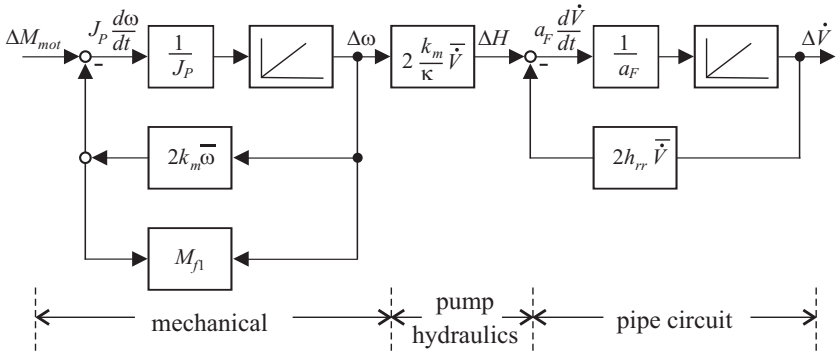


Fig. 6.5. Signal flow for the linearized behavior of a pump–pipe system in a closed circuit around the steady state $\left[\bar{\omega}, \bar{\dot{V}}\right]$ for a small change of the motor torque

Identifying the static pump characteristic by variation of the valve position, speed and flow are no longer proportional to each other. Therefore, the complete centrifugal pump model has to be used. For the head characteristic it holds that

$$H = h_{nn}\omega^2 - h_{nv}\omega\dot{V} - h_{vv}\dot{V}^2 \tag{6.1.22}$$

and for the torque

$$M_P = \rho g \left(h_{nn}\omega\dot{V} - h_{nv}\dot{V}^2 - h_{vv}\frac{\dot{V}^3}{\omega} \right) \tag{6.1.23}$$

Experiments have shown, [6.22], [6.24], that the torque for the investigated pump can be better approximated by

$$M_P = k_0\omega\dot{V} - k_1\dot{V}^2 + k_2\omega^2 \tag{6.1.24}$$

By changing the valve position with constant motor speed the head H, the motor torque M and the volume flow $\dot{V}$ are varied along the characteristic curves. Least-squares parameter estimation is then based on

$$H = \boldsymbol{\Psi}^T\boldsymbol{\Theta}_H \tag{6.1.25}$$

$$M_P = \boldsymbol{\Psi}^T\boldsymbol{\Theta}_M \tag{6.1.26}$$

with the data vector

$$\Psi^T = \left[\omega^2 \;\; \omega\dot{V} \;\; \dot{V}^2\right] \tag{6.1.27}$$

and the parameter vector

$$\boldsymbol{\Theta}_H^T = [h_{nn} \quad -h_{nv} \quad -h_{vv}] \tag{6.1.28}$$

$$\boldsymbol{\Theta}_M^T = [k_2 \;\; k_0 \quad -k_1]\,. \tag{6.1.29}$$

The experiments were performed for a volume flow range of $0 - -150\%$ relative to the rated flow. However, a range of $60 - -150\%$ is sufficient for fault detection.

The shut-off experiments leads to a jump of the driving torque to zero. If it is assumed that the pump speed and the flow are proportional to each other because the flow resistance remains constant, the terms $\omega\dot{V}$ and $\dot{V}^2$ can be replaced by ω^2, see (6.1.15) and (6.1.17). This yields with (6.1.7), neglecting the friction terms,

$$M_{mot}(t) = J_P\, d\omega/dt + k_m\omega^2 \tag{6.1.30}$$

Note that both parameters J_P and k_m can be identified, although $M(t)$ is zero except for $t \leq 0$. The reason is that for identification the values $M(t)$ and $\omega(t)$ are not used directly but are first sent through a (low-pass) state-variable filter in order to obtain the derivative $d\omega/dt$. Therefore, the filtered value of $M(t)$ is not a jump to zero but a slow transition to zero which provides sufficient excitation for parameter identification. The term $k_m\omega^2$ represents the load at shut-off. Generally, an increase of load means a decrease of the shut-off period.

The centrifugal pump model was now used for incipient fault detection. The idea was to identify the slowly proceeding deterioration of the pump state due to abrasive wear, erosion and cavitation by identifying the parameters of the pump model. Several faults were built into the pump to test the reaction of the fault-detection method:

1) *Wear at the clearance gap*. The clearance gap ring is one of the most frequently changed spare parts of a centrifugal pump. A widening of the clearance gap increases the leakage flow from impeller outer diameter to suction side and significantly reduces pump efficiency, especially with pumps of low specific speed. However, the degree of wear depends very much on the operating conditions of the pump, especially the quality of the fluid medium (concentration of salt, sand, ashes, etc.) and the operating point (pump delivery flow, available net pressure suction head (NPSH)). Thus, no universal recommendation can be given for when to change the clearance ring. Several experiments with different rings were carried out. The effect of an increased clearance gap is a leftward shift of the characteristic curves which can be seen in Figure 6.6. Furthermore, the disk friction increases for single-stage pumps, because the clearance gap flow enters the back part of the impeller with a peripheral velocity, which gets faster as the flow passes from the outer diameter to the inner diameter. The shut-off period becomes smaller (especially with shut-off at zero flow), because the flow through the clearance represents an additional load, see Figures 6.7 and 6.8. Thus, the symptoms for clearance gap wear are, see Table 6.3, a reduction of h_{nn} and an increase of h_{vv}, k_2 and k_m.
2) *Wear at impeller outlet*. Experiments with different impeller faults were performed, out of which the wear at the impeller outlet is further discussed here. Abrasive wear can lead to erosion at the pressure side of the blades. This has a similar effect as a sharpening of the blades: The decrease of blade width at the outlet has a positive effect on the flow conditions and thus first increases the head of the pump. Therefore, an increased pump head is not always a sign of an improved pump state. The effect on the parameters is mainly an increase in h_{nn}, h_{nv} and k_2.
3) *Deposits at impeller inlet*. These can severely hamper the operation of a pump and in extreme cases can lead to total pump failure. In many cases these deposits consist of organic waste but there may also occur a slow sedimentation by chemical reactions, especially if the pump has a standstill. Therefore, an identification of this fault is interesting for the periodic check of standby pumps, especially in chemical industry. A partial obstruction of the pump inlet manifests itself mainly at nominal load or overload as can be seen in Figure 6.9. During the experiments there were first a few threads fastened at the entrance of the blades and then a lot of threads. In both cases, a determined deviation of the pump‘s characteristic curve could be observed. There was a pronounced increase of the parameter h_{vv} and a decrease of h_{nn}.

Statistical evaluation of the resulting parameter behavior was done to distinguish between normal random parameter variations and significant variations in the case of a fault. Table 6.3 summarized the results of some experiments. It can be seen that

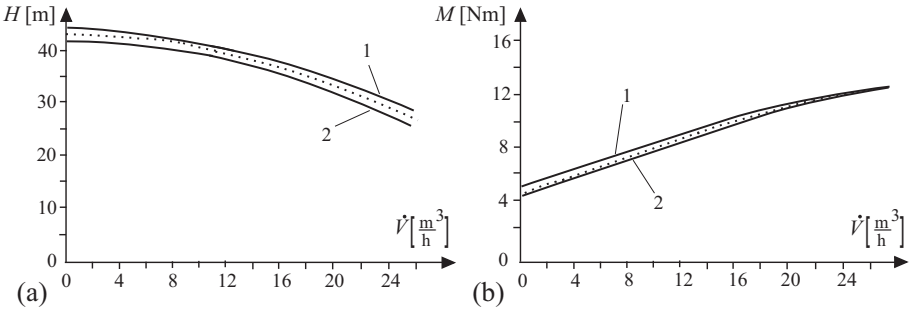


Fig. 6.6. Pump characteristics as a function of wear at the clearance gap: a) head-flow characteristic. Dotted line: reference characteristic; b) torque-flow characteristic. Dotted line: reference characteristic. 1 very small clearance gap; 2 large clearance gap

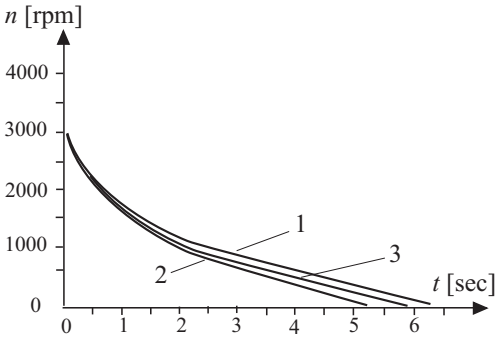


Fig. 6.7. Speed for pump shut-off in the case of different clearances. 1 very small clearance; 2 large clearance gap; 3 reference gap

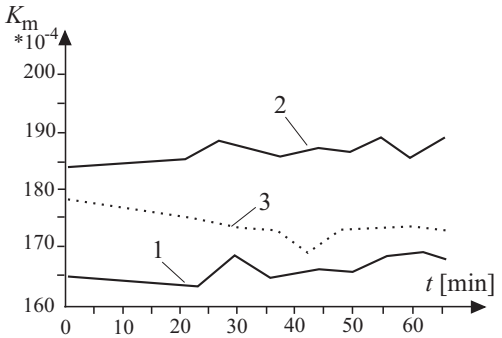


Fig. 6.8. Estimated parameter k_m as a function of clearance gap. Dotted line: reference characteristic. 1 very small clearance; 2 large clearance

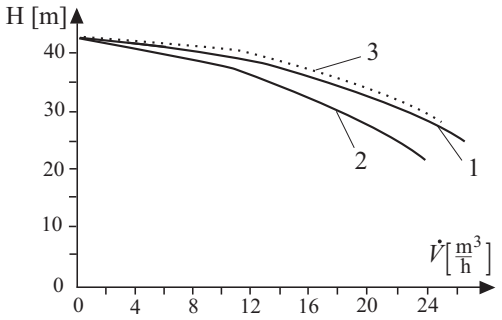


Fig. 6.9. Influence of deposits at impeller inlet on the head-flow characteristic. Dotted line: reference characteristic. 1 small deposits; 2 large deposits

Table 6.3. Symptoms from parameter estimation of static characteristics and shut-off dynamics

	Static head characteristic			Static torque characteristic			Shut-off dynamics	
Fault	h_{nn}	h_{nv}	h_{vv}	k_0	k_1	k_2	J_P	k_m
wear at clearance gap	−−	−	++	0	+	++	0	++
small deposits at impeller outlet	−	0	++	+	+	0	0	0
deposits at impeller inlet	−	++	++	0	0	+	0	0
abrasive wear at impeller outlet	+	+	0	0	0	+	0	0
broken blade	+	0	−	0	−	+	+	+
caviation at impeller inlet	−	−	0	0	0	−	0	+

the main hydraulic damage due to abrasive wear and erosion can clearly be detected. If the size of the symptoms are taken into account the faults can be isolated. Some more practical results for pump diagnosis are given in [6.24]. These investigations have shown that a constant speed pump can be diagnosed by parameter estimation of characteristics with changing valve position and by shut-off experiments. As erosion damage is of a slowly proceeding nature which usually does not change the pump operation significantly, model-based pump diagnosis with parameter estimation provides an effective method for the implementation of a predictive maintenance schedule.

The application of parameter estimation for the fault detection of centrifugal pumps was also investigated in [6.8]. By using the basic static pump equation (6.1.3) three parameters were estimated for different positions of a flow valve. Remaining small errors in the resulting delivery head are then approximated with a multilayer perceptron neuronal net. Deviations of the measured head $\Delta H(n, \dot{V}, T) = H_{meas} - H_{mod}$ are then used to estimate deviation parameters $\Delta\hat{\Theta}$. The deviations of these three parameters of (6.1.28) then allow one to distinguish between rotor faults

like deposits, broken blade, cavitation erosion and clearance gap wear. The influence of the fluid temperature on the model accuracy is observable for $20°C \leq T_{fl} \leq 80°C$ and is, however, negligible for the final fault classification. Hence, similar results were obtained as described above.

b) Dynamic operation for stepwise speed changes

The centrifugal pump is now driven by a speed-controlled DC motor and pumps water through a closed pipe circuit, see Figure 6.4. Both, the DC motor and the pump are now considered as a unit, [6.6].

The measured signals are: U_2, armature voltage; I_2, armature current; $\dot{V}$, volume flow rate; ω, angular velocity; H, pump total head.

The basic equations after some simplifying assumptions are:

a) armature circuit

$$L_2 \frac{dI_2(t)}{dt} = -R_2 I_2(t) - \Psi\omega(t) + U_2(t) \tag{6.1.31}$$

b) mechanics of motor and pump

$$J_P \frac{d\omega}{dt} = \Psi I_2(t) - M_{f0} - \rho g h_{th1} \omega(t) \dot{V}(t) \tag{6.1.32}$$

c) hydraulics of the pump, [6.26]

$$H(t) = h_{nn}\omega^2(t) - h_{nv}\omega(t)\dot{V}(t) - h_{vv}\dot{V}^2(t) = h'_{nn}\dot{V}^2(t) \tag{6.1.33}$$

In this case all three terms can again be lumped together, as $\dot{V}$ is proportional to ω, see (6.1.15).

d) hydraulics of the pipe

$$a_F \frac{d\dot{V}(t)}{dt} = -h_{rr}\dot{V}^2(t) + H(t) \tag{6.1.34}$$

The overall model is basically nonlinear but linear in the parameters to be estimated. Therefore least-squares parameter estimation can be applied in its direct, explicit form described in Section 2.5.2. The models contain nine process coefficients:

$$\mathbf{p}^T = \left[L_2, R_2, \Psi, J_P, M_{f0}, h_{th1}, h'_{nn}, a_F, h_{rr}\right]. \tag{6.1.35}$$

For the parameter estimation the equations are brought into the form

$$y_j(t) = \Psi_j^T(t)\hat{\boldsymbol{\Theta}}_j, \; j = 1, 2, 3, 4 \tag{6.1.36}$$

where

$$\left.\begin{array}{ll} y_1(t) = dI_2(t)/dt & y_2(t) = d\omega(t)/dt \\ y_3(t) = H(t) & y_4(t) = d\dot{V}(t)/dt \end{array}\right\} \tag{6.1.37}$$

The model parameters

$$\hat{\boldsymbol{\Theta}}^T = \left[\hat{\boldsymbol{\Theta}}_1^T, \hat{\boldsymbol{\Theta}}_2^T, \hat{\boldsymbol{\Theta}}_3^T, \hat{\boldsymbol{\Theta}}_4^T\right] \tag{6.1.38}$$

were estimated by the least-squares method in the form of discrete square-root filtering (DSFI). Based on the model parameter estimates $\hat{\boldsymbol{\Theta}}$ all nine process coefficients of $\mathbf{p}$ could be calculated uniquely.

The DC motor is controlled by an AC/DC converter with cascade control of the speed and the armature current as auxiliary control variable. The manipulated variable is the armature current U_2. A microcomputer DEC-LSI 11/23 was connected online to the process. For the experiments the reference value $W(t)$ of the speed control has been changed stepwise with a magnitude of 750 rpm every 2 min. The operating point was $n = 1000\,\text{rpm}$, $H = 5.4\,\text{m}$ and $\dot{V} = 6.48\,\text{m}^3/\text{h}$. The signals were sampled with sampling time $T_0 = 5\,\text{ms}$ and 20 ms over a period of 2.5 and 10 s, so that 500 samplings were obtained. These measurements were stored in the core memory before estimation. Hence, one set of parameters and process coefficients was obtained every 120 s. For the training phase 50 coefficient sets were used. Table 6.4 gives an overview of significant changes of process coefficients for 19 different artificially generated faults. A selection of experiments will now be considered in more detail.

1) *Fault A5*: *Disturbance of the air cooling of the DC motor*, Figure 6.10
   A stepwise reduction (20%, 50%, 75%, 100%) of the air flow (e.g. due to plugging by dirt) leads to a temperature change of the whole motor and therefore to an increase of the resistance in the armature circuit and excitation circuit. Therefore, R_2 increases and the magnetic flux linkage Ψ decreases. This is an example where the coefficients move in opposite directions.
2) *Fault P3*: *Increase of slot clearance of the centrifugal pump*, Figure 6.11
   An increase of the slot clearance between the pump wheel and the pump case increases the internal losses. Therefore, h_{th1} increases and h'_{nn} decreases.
3) *Fault F1a*: *Cavitation in the centrifugal pump*, Figure 6.12
   A small cavitation and gas bubble generation in the pump by lowering the entrance pressure is indicated by an increasing coefficient a_F, which is proportional to the time constant of the pipe system, see (6.1.13).

These experiments have demonstrated that in all cases, where a significant change of process coefficients could be expected the fault could be detected. Based on the patterns given in Table 6.4 most of the faults can be isolated. At least motor faults are clearly isolable from pump faults.

6.1.4 Fault detection with nonlinear parity equations and parameter estimation

In order to develop an online fault detection and diagnosis method for a centrifugal pump–pipe–tank system over a large operating range, a plant according to Figures 6.13 and 6.14 is considered. The pump is driven by an inverter-fed, speed-variable induction (squirrel-cage) motor which is speed-controlled by a field-oriented controller. The stator current vector $\mathbf{I}_s = I_{s\alpha} + i\,I_{s\beta}$ is measured and transformed in the

Table 6.4. Detected symptoms for the DC motor and centrifugal pump based on parameter estimation for stepwise speed changes. +: positive change, −: negative change, 0: no change

Fault	Symptom								
	L_2	R_2	Ψ	J_P	M_{f0}	h_{th1}	h'_{nn}	a_F	h_{rr}
A1: exc. resistance increase	0	0	−	0	0	0	0	0	0
A2: armature resistance increase	0	+	−	0	0	0	0	0	0
A3: affected brushes	−	+	−	0	0	0	0	0	0
A4: new brushes	0	0	0	0	0	0	0	0	0
A5: insufficient cooling	0	+	−	0	0	0	0	0	0
A6: cold drive	0	0	+	−	0	0	0	0	0
K1: shaft displacement	0	0	0	−	0	0	0	0	0
P1a: bearing without grease	0	0	0	0	−	0	0	0	0
P1b: bearing with dirt	0	0	0	+	+	+	0	0	0
P2: side thrust compensation defective	0	0	0	0	0	0	0	0	0
P3: splitting clearance increase	0	0	0	0	0	+	−	0	0
P4: affected impeller	0	0	0	−	0	+	−	0	0
P5a: pump casing defective I	0	0	0	0	+	0	0	0	0
P5b: pump casing defective II	0	0	0	0	−	+	0	+	0
F1a: little cavitation	0	0	0	0	0	0	0	+	0
F1b: medium cavitation	0	0	0	0	0	+	−	+	−
F2: insufficient venting	0	0	0	0	0	+	−	−	−
F3: fluid temperature increase	0	0	0	0	0	0	−	0	−
F4: valve position increase	0	0	0	0	0	+	0	0	+

reference frame defined by the rotor flux $\mathbf{I}_s = I_{sd} + i\,I_{sq}$ which is obtained by using an adequate model, see Section 3.2.

The motor torque can then be determined by

$$M_{mot} = k_T \Psi_R I_{sq} \tag{6.1.39}$$

where k_T is known from the motor data sheet.

Further measurements are:

p_1 pump pressure inlet
p_2 pump pressure outlet
ω pump speed
$\dot{V}$ volume flow.

Figure 6.15 shows the overall configuration of the investigated pump–pipe system. The mathematical models of the pump used have to be adapted to the pump–pipe system, [6.5], [6.30]. Based on the theoretically derived equations from Section 6.1.2 the following models are used here:

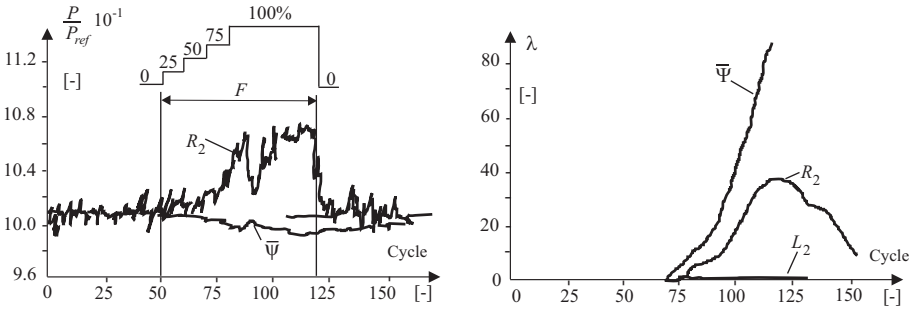


Fig. 6.10. Change of process coefficients for reduction of the cooling air flow. (λ) is the test quantity of a Bayes decision test, [6.6]

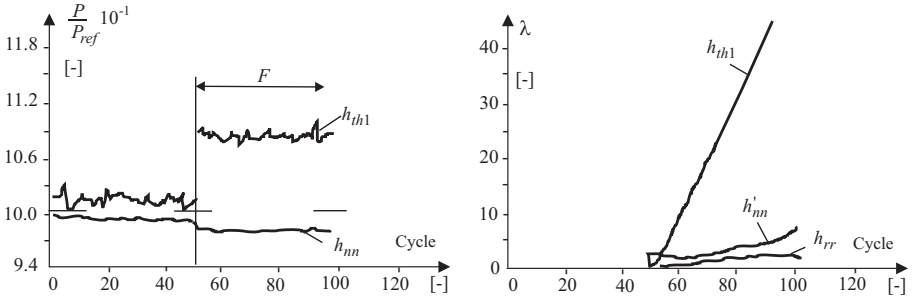


Fig. 6.11. Change of process coefficients after increasing the slot clearance of the pump

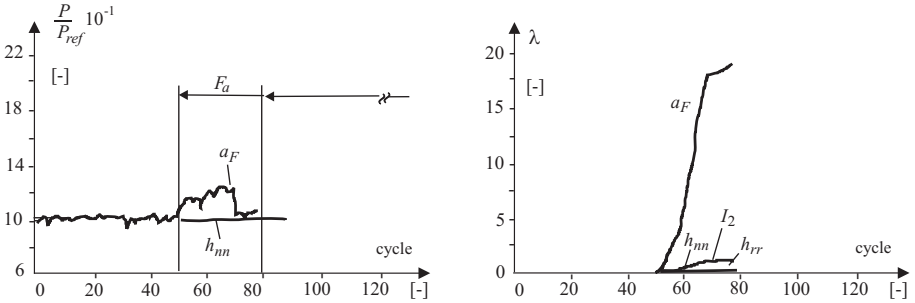


Fig. 6.12. Change of process coefficients in the case of small cavitation and gas bubble generation in the pump

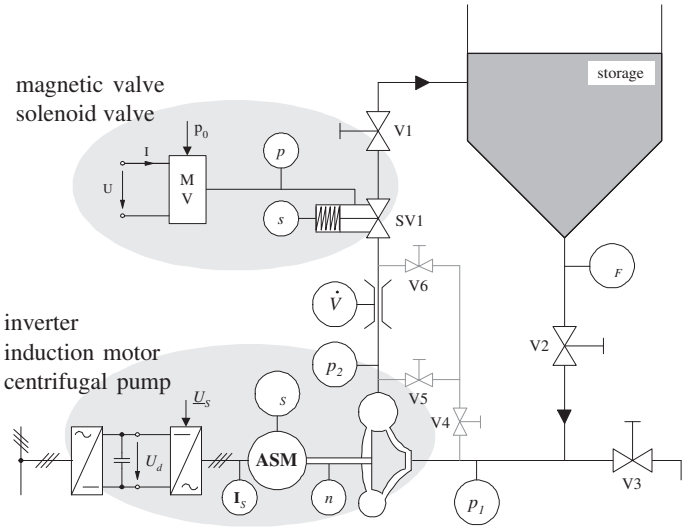


Fig. 6.13. Centrifugal pump–pipe–tank plant with measurements: AC motor: Siemens 1 LA 5090-2AA (norm motor) $P_N = 1.5\,\text{kW}$; $n_W = 2900\,\text{rpm}$; frequency converter: Lust MC 7404; circular pump: Hilge, $H_{max} = 130\,\text{m}$; $\dot{V}_{max} = 14\,\text{m}^3/\text{h}$; $P_{max} = 5.5\,\text{kW}$

Fig. 6.14. Photo of the investigated centrifugal pump–pipe–tank plant

$$H(t) = h_{nn}\omega^2(t) - h_{nv}\omega(t)\dot{V}(t) - h_{vv}\dot{V}^2(t) \tag{6.1.40}$$

$$H(t) = \frac{p_2(t) - p_1(t)}{\rho g} = \frac{\Delta p(t)}{\rho g} \tag{6.1.41}$$

$$H(t) = a_F d\dot{V}(t)/dt + h_{rr}\dot{V}^2(t) \tag{6.1.42}$$

$$J_p\dot{\omega}(t) = M_{mot}(t) - M_{th}(t) - M_f(t) \tag{6.1.43}$$

$$M_{th}(t) = M_{th1}\omega(t)\dot{V}(t) - M_{th2}\dot{V}^2(t) \tag{6.1.44}$$

$$M_f(t) = M_{f0}\ \text{sign}\ \omega(t) + M_{f1}\omega(t) \tag{6.1.45}$$

A comparison of these theoretically derived equations has shown that because of (6.1.11), neglect of the viscous friction in (6.1.32) and (6.1.8), the following simplified relations can be used, [6.30]:

$$\Delta p(t) = \tilde{h}_{nn}\omega^2(t) - \tilde{h}_{\omega}\omega(t) \tag{6.1.46}$$

$$J_P\dot{\omega}(t) = M_{mot}(t) - M_{f0}(t) - M_2\omega^2(t) \tag{6.1.47}$$

These models agree also with a larger pump–pipe system, [6.5]. Figure 6.16 shows the resulting signal flow diagram.

a) Measurement of $I, \omega, \Delta p, \dot{V}$

Based on these models and after discretizing, the following residuals can be obtained, compare Figure 6.17 and [6.30], [6.32]:

Static pump model (6.1.46):

$$r_1(k) = \Delta p(k) - w_1\omega^2(k) + w_2\omega(k) \tag{6.1.48}$$

Dynamic pipe model (6.1.42):

$$r_2(k) = \dot{V}(k) - w_3 - w_4\sqrt{\Delta\hat{p}(k)} - w_5\dot{V}(k-1) \tag{6.1.49}$$

Dynamic pump–pipe model (6.1.42), (6.1.46)

$$r_3(k) = \dot{V}(k) - w_3 - w_4\sqrt{\Delta\hat{p}(k)} - w_5\dot{V}(k-1) \tag{6.1.50}$$

$$\Delta\hat{p}(k) = w_1\omega^2(k) - w_2\omega(k) \tag{6.1.51}$$

Dynamic inverse pump model (6.1.48)

$$r_4(k) = M_{mot}(k) - w_6 - w_7\omega(k) - w_8\omega(k-1) - w_9\omega^2(k) - w_{10}M_{el}(k-1) \tag{6.1.52}$$

The residuals $r_1(k)$), $r_2(k)$ and $r_3(k)$ are output residuals which follow by comparing the measured $\Delta p(k)$ and $\dot{V}(k)$ with the corresponding model outputs, see Figure 6.17. However, $r_4(k)$ is an input residual, because $M_{mot}(k)$ is compared with the output of an inverse pump model. $r_2(k)$ and $r_3(k)$ include flow sensor dynamics of first order. The sampling time is $T_0 = 10$ ms.

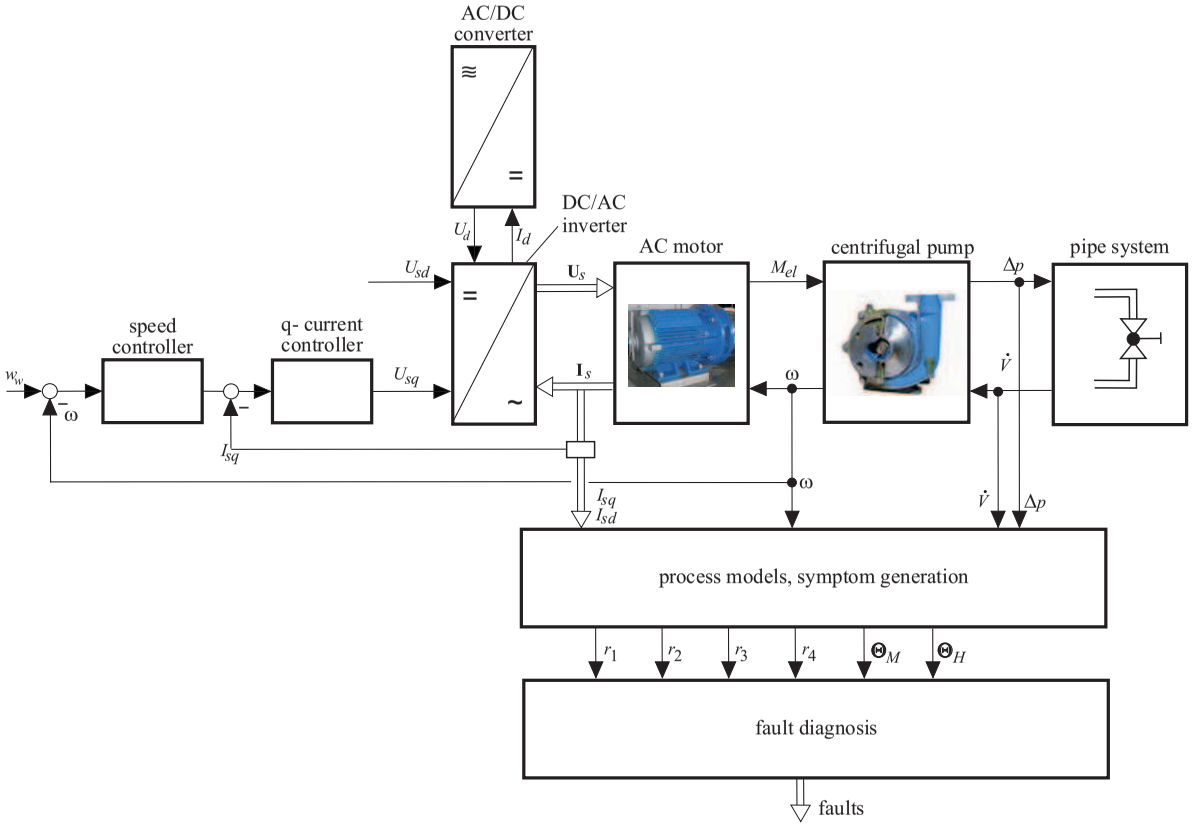


Fig. 6.15. Scheme of the model-based fault detection and diagnosis system for a speed-controlled centrifugal pump with parity equations and parameter estimation. $\mathbf{U}_s$: pulse-width-modulated stator voltage vector; $\mathbf{I}_s$: stator current vector

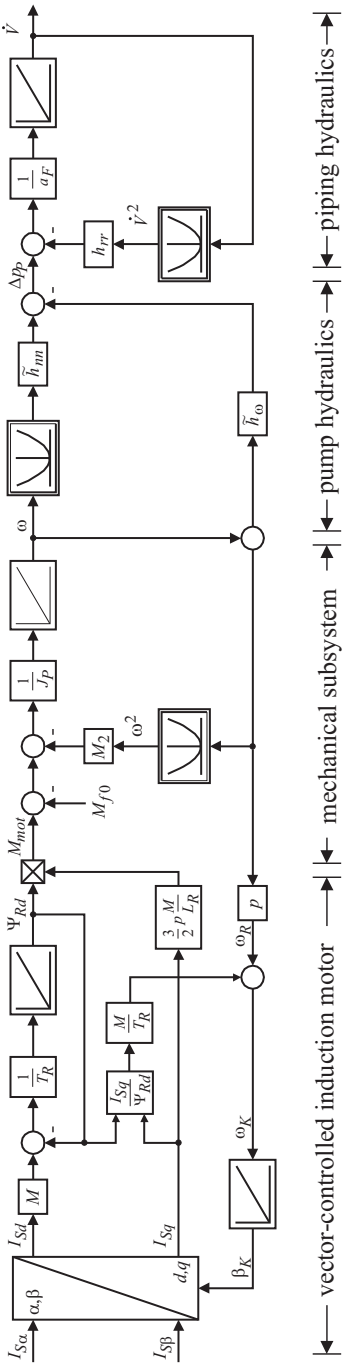


Fig. 6.16. Signal flow diagram of the simplified nonlinear overall motor–pump–pipe system

The parameters $w_1, \ldots, w_{10}$ follow directly from known physical data described in the equations above or are estimated, e.g. with methods of least squares based on measurements of $I_{sq}(t)$, $\omega(t)$, $\Delta p(t)$ and $\dot{V}(t)$. However, the parameters w_i depend, especially for low speed at the operating point. Therefore, for each residual a multi-model approach is used. It has turned out that it is sufficient to consider the parameters dependent on the angular speed only.

The pump system was excited by changing the speed according to an amplitude-modulated PRBS over the whole operating range, and with the local linear model network LOLIMOT the parameters $w_i(\omega), i = 1, \ldots, 10$ were determined using three local models each, see [6.15]. Figure 6.18 shows a comparison of measured and reconstructed values with the models. Hence, a very good agreement can be stated.

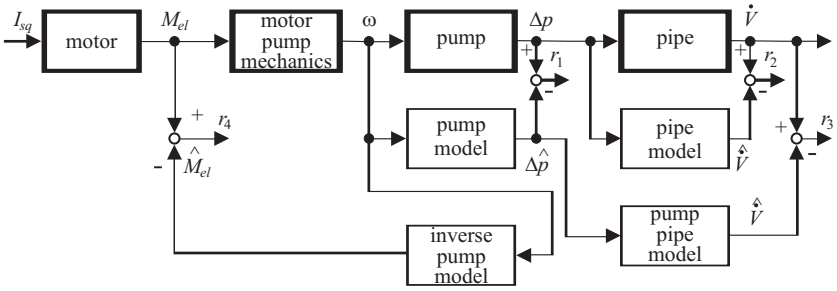


Fig. 6.17. Residual generation with parity equations for the pump–pipe system

The following faults were introduced into the pump–pipe system:

- Offset sensor faults ω, $\dot{V}$, p_1, p_2
- Increased resistance by piecewise closing of a valve after the pump
- Cavitation by piecewise closing of a valve before the pump
- Increased bearing friction by removing grease and introducing iron deposits
- Impeller defect by closing one channel between two vanes with silicon
- Sealing-gap losses by opening a bypass valve
- Leakage between pump and flow measurement.

Table 6.5 shows the resulting symptoms. The residuals of the parity equations can be obtained without input excitation, i.e. in steady-state operation. The resulting residuals indicate that the sensor offset faults, sealing-gap losses and increased bearing friction are strongly isolable. However, increased flow resistance, cavitation and impeller defect are either only weakly or not isolable. This means that all the faults are detectable but some of them cannot be differentiated. As in the case of dynamic excitation the nonlinear models are very precise, passing deviations of the residuals result. In order to avoid over-large thresholds, adaptive thresholds are used. In addition to a constant value, the thresholds depend on a high-pass-filtered value of the speed ω, which increases the threshold in the case of a speed change, [6.30], [6.31].

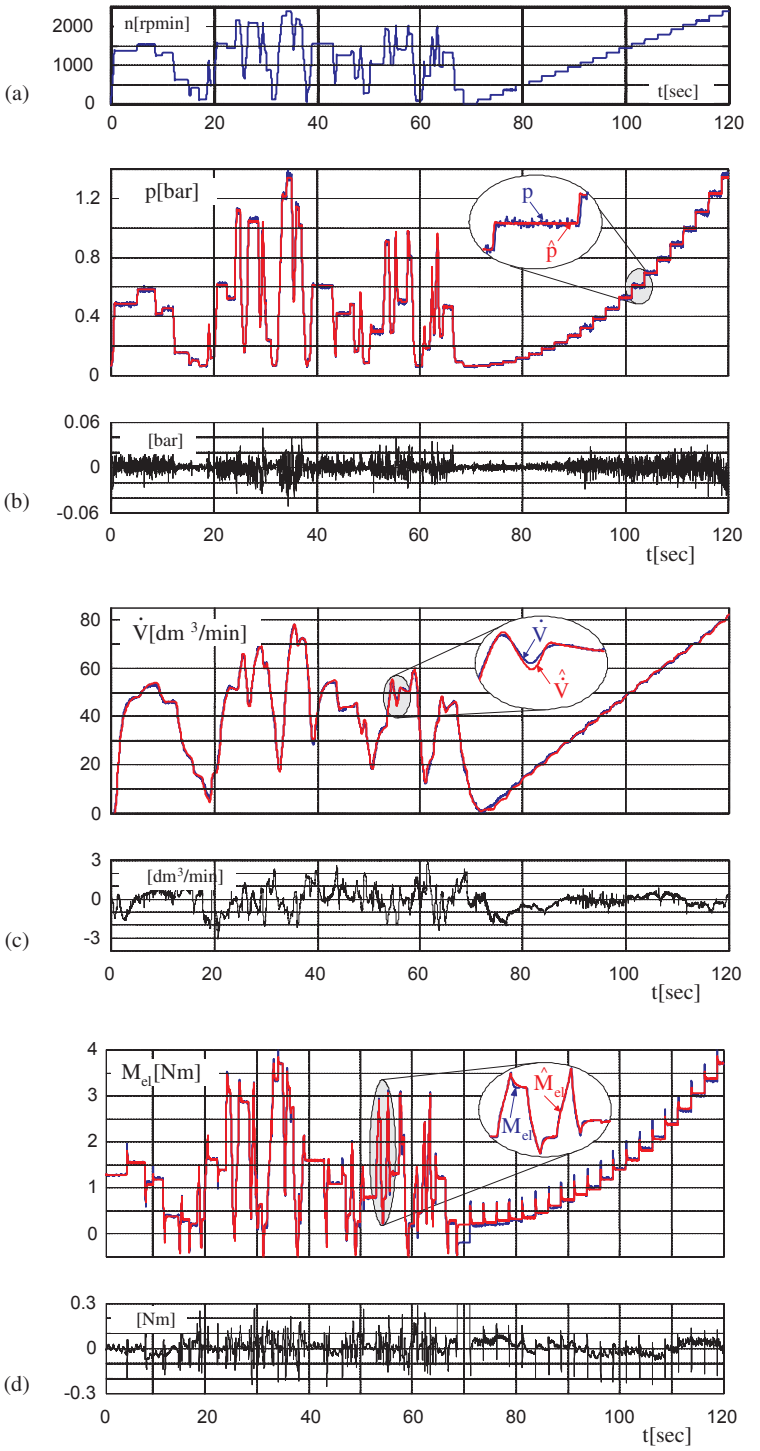


Fig. 6.18. Measured signals of the pump–pipe system, LOLIMOT model outputs and their differences: a) angular speed; b) delivery pressure difference; c) flow rate; d) torque of AC motor

A dynamic excitation with a test PRBS for the speed also allows one to estimate the parameters of the models (6.1.42), (6.1.46) and (6.1.47) with a recursive least-squares method as described in Section 2.5, see also [6.13], Section 2.5. The changes of these physically defined parameters are given in Table 6.5 and show that now all faults are isolable and can therefore be diagnosed by combining the two methods.

b) Measurement of I, ω

If the delivery pressure Δp and the flow rate $\dot{V}$ are not measurable, the residual r_4 can be calculated based on measured speed ω and motor current I_{sq}. This allows one to detect a sensor fault in ω and some pump faults. Additional parameter estimation enables one to determine parameter deviations of J_P, M_{f0} and M_2 with (6.1.47) and to isolate some more pump faults.

Similar results as described above have been obtained by [6.5] for a larger pump with $P = 3.3\,\text{kW}$ and $\dot{V}_{max} = 150\,\text{m}^3/\text{h}$ and a larger pipe circulation system with two heat exchangers. Two different flow meters could be used. This allowed the generation of six residuals and four parameter estimates. Together with two variances of residuals, altogether 13 symptoms could be obtained, which enabled the diagnosis of 11 different faults of sensors, pump and pipe system.

These symptoms were then used to train 20 fuzzy rules with the SELECT procedure described in [6.15], Section 17.3.5, yielding a 100% classification accuracy.

Table 6.6 enables one to see which faults are only detectable or also diagnosable with combined parity equations and parameter estimation. A minimal measurement of the torque $M = f(I)$ and speed ω allows one to detect some few faults but not to diagnose them. By adding a sensor for p_1 and p_2, or for Δp, many more faults can be detected and diagnosed. The additional implementation of a flow rate sensor has little influence on the number of detectable faults, but allows one to diagnose many more faults. This shows that model-based detection of faults is possible with three to four sensors, but that the fault diagnosis is improved considerably by one additional sensor (here the flow rate).

6.1.5 Fault detection with vibration sensors

Rotating machinery such as centrifugal and reciprocating pumps generates certain oscillations. In the case of centrifugal pumps these oscillations are generated by the rotating shafts, blades of the rotors, ball bearings, unbalance and fluid oscillations through turbulence and vortex or cavitation. The arising frequencies therefore depend on the rotational speed, flow rate and special fluid-phenomena.

Therefore is is obvious to detect changes, malfunctions and faults through vibration and oscillation measurements, especially also as some faults can be detected by the human ear. The available sensors are, e.g. oscillation velocity sensors with a low-frequency suspended seismic mass and inductively generated voltage or oscillation accelerometers with a high-frequency suspended seismic mass or with piezoelectrical sensors possessing small mass and large spring stiffness.

Table 6.5. Fault-symptom table for the pump–pipe system obtained by parity equation and parameter estimation: 0: small values; +/−: positive/negative deflection; + + / − −: strong positive/negative deflection

	Symptoms										
	parity equations				parameter estimates						
Fault	$\|r_1\|$	$\|r_2\|$	$\|r_3\|$	$\|r_4\|$	ΔJ_P	ΔM_{f0}	ΔM_2	Δa_F	Δh_{rr}	$\Delta \tilde{h}_{nn}$	$\Delta \tilde{h}_{\omega}$
Sensor ω	+/++	0	+/++	+/++	0	0	0	0	0	0/−	+/++
Sensor $\dot{V}$	0	+/++	+/++	0	0	0	0	0/−	+	0	0
Sensor p_1, p_2	+/++	+/++	0	0	0	0	0	0/+	0	0/−	+/++
Gap losses	+	0	0	0	0	0	0	0	0	−	+
Leakage	+	+	+	0	0	0	0	−	−	0	0
Incr. flow resist. 20–40%	+	+	+	0	0	0	0	0	+	0	0
Incr. flow resist. 40–60%	+	++	+	+	0	0	0	++	++	0	0
Incr. flow resist. 60–90%	++	++	++	++	0	0	+	++	++	0	0
Cavitation	+	++	++	+	0	0	+	−−	++	−−	++
Incr. bearing friction	0	0	0	+	0	+	+	0	0	0	0
Impeller defect	+	+	+	+	+	0	0	−	−	−	+

Table 6.6. Detectable and diagnosable faults in dependence on the sensors used. Assumed is a modern frequency converter that is able to reconstruct the motor torque M without additional sensors. Omitted are faults at the frequency converter itself and other electric faults in the motor. Faults in parentheses are difficult to identify yet not completely impossible (depends on the individual setup), [6.5].

	Sensor usage				
Fault detectable	M	M, ω	M, ω, p_2	M, ω, p_1, p_2	$M, \omega, p_1, p_2, \dot{V}$
Total breakdown	x	x	x	x	x
Defective blade wheel		(x)	x	x	x
Incr. shaft or motor friction			x	x	x
Sensor fault ω		x	x	x	x
Sensor fault $\dot{V}$					x
Sensor fault p_1, p_2			p_2	x	x
Decreased flow resistance		(x)	(x)	x	x
Increased flow resistance		(x)	(x)	x	x
Cavitation through pressure reduction		(x)	(x)	x	x
Insufficient de-ventilation of sensors p_1, p_2			p_2	x	x
Insufficient de-ventilation of sensor $\dot{V}$					x

	Sensor usage				
Fault diagnosable	M	M, ω	M, ω, p_2	M, ω, p_1, p_2	$M, \omega, p_1, p_2, \dot{V}$
Total breakdown		x	x	x	x
Defective blade wheel			x	x	x
Incr. shaft or motor friction			x	x	x
Sensor fault ω					x
Sensor fault $\dot{V}$					
Sensor fault p_1, p_2					x
Decreased flow resistance					x
Increased flow resistance					x
Cavitation through pressure reduction					x
Insufficient de-ventilation of sensors p_1, p_2			p_2	x	x
Insufficient de-ventilation of sensor $\dot{V}$					x

During the last two decades several investigations were reported for fault detection with vibration sensors.

The application of four piezoelectric acceleration sensors in combination with two pressure and one motor current phase Hall sensor for fault detection was investigated in [6.20]. Sampling frequency was 125 kHz over a time interval of 1.4 s or 70 pump rotations with constant speed (3000 rpm). The evaluation of the structure-borne noise was made in the spectral ranges 5–10 kHz, 10–15 kHz and 15–20 kHz. The passing of the rotor blades results in a basic frequency in the pressure spectra and is influenced by faults like cavitation and broken blades. Several pump faults could be differentiated with classification methods and each one acceleration sensor for the pressure and for the pump casing together with the motor phase current seem to be sufficient.

[6.17] investigated pressure and casing fluctuations by using signal analysis together with neuro-fuzzy methods. Fourier analysis of two orthogonal vibration sensors to detect axle motions relative to sliding bearings indicated unbalance, blocking and wear.

A series of publications on the analysis of structure-borne noise with different kinds of acceleration sensors at pump casings or integrated in the containment shell of magnet coupling pumps are [6.16], [6.10], [6.19], [6.12], [6.9], [6.25]. They show the application of different signal-analysis methods, the generation of features and computerized learning methods. Several faults could be detected and classified for a definite pump speed. However, the application for variable-speed pumps and the transferability to other pump sizes and pump types seem not to be directly possible. Also finding the best locations for positioning the accelerometers is important.

[6.11] reports on the frequency analysis of pressure and flow rate and shows how the frequency spectra are changed by damage to the impeller, cavitation, two-phase flow and restrictions.

A high-frequency analysis of the motor current allows one to gain information on stochastic and periodic disturbances of the torque, see also Section 3.2.2. An application for underwater well pumps has shown that the current frequency spectrum between 5 and 100 Hz allows one to detect increased bearing gap, recirculation effects at low volume flows and congestions, [6.21].

The advantage of the vibration sensors is that they can be applied easily at the pump casing, if an appropriate position has been found. However, the vibrations generated by the rotor and by the flowing medium are proportional to the square of the speed and the flow rate. Therefore the amplitudes of the signals are relatively small for lower speeds and part load and tend not to give useful signals. In addition, the analysis of the signals with FFT, parametric ARMA models or wavelets needs high-frequency sampling and computing power. It is further subject to noise and vibrations from other plant equipment. Vibration analysis of pumps is in principle limited to the detection of faults which generate oscillations through the rotation, flow and special fluid phenomena, like cavitation. Because many other faults cannot be detected, it should be combined with other detection methods, if other measurements such as Δp, ω, M, $\dot{V}$ are available. If, however, the pumps are not well instrumented and run with constant speed, vibration analysis is a first choice.

6.1.6 Conclusions

The described methods for the model-based fault detection and diagnosis of centrifugal pumps are based on *standard measurements* such as voltage, current and angular speed of the electrical drive and inlet and outlet pressure or difference pressure (head) and volume or mass flow rate of the fluid. It has been shown that fault detection as well as fault diagnosis is improved considerably by using the physical relations between the measured variables in the form of process models and that an in-depth diagnosis is increasingly possible the more sensors are available. A considerable advantage of the described process-model-based fault-detection methods is that they apply physically derived pump models. Therefore they are applicable to a wide range of operating points and they are directly transferable to other centrifugal pumps. Further their symptoms are usually well interpretable and understandable.

The applicability of the different methods depends on the *kind of operation*, as summarized in Table 6.7. If the pump is operating for long time periods in *one steady state*, i.e. with constant speed and flow, then mainly parity equations with various residuals of input and output variables, as shown in Figure 6.17 and Table 6.5 can be used. The number of diagnosable faults increases with the number of applied sensors, Table 6.6. The situation improves if the pump is operating at *different steady states*, i.e. speed and volume change from time to time. Then physical- and/or experimental-based algebraic equations for the pump characteristics can be taken for steady-state parameter estimation of some coefficients which express fluid dynamic phenomena and different losses. The changes of these coefficients can be traced back to different pump faults, see Figures 6.6 and 6.12 and Table 6.3, and allow one to diagnose some faults because of different influences on the coefficients (signs and sizes). The sign of the changes may depend on the pump type and operating point, [6.22]. The structure of the algebraic characteristic equations is in principle based on theoretical pump models, however may need some simplification or adaption for the individual pump and pipe system, also depending on operation in an open or closed fluid flow circuit.

If the pump is driven continuously or from time to time in *dynamic (transient) states* by fast changes of the electrical motor speed or the fluid flow by valves, then dynamic pump and pipe equations can used for parameter estimation, as depicted in Figures 6.10 to 6.12 and Table 6.4. Additionally, parity equations can be applied, resulting in some residuals. Table 6.5 shows that the combination of these two fault-detection methods allows one to generate a large number of symptoms and to differentiate between several sensors and pump faults.

A further possibility for fault detection is to observe the speed behavior of the pump after *shut-off* of the electrical drive. Figure 6.7 and Table 6.3 show that some pump faults can be detected. However, this depends on the size of the connected pipe system. It may increase the number of generated symptoms, especially for short pipe lengths.

Note that all the model-based methods allow one to detect faults in the motor as well as in the connected pipeline in addition to those in the pump. The described model-based methods of fault diagnosis can be supplemented by *signal-analysis-*

Table 6.7. Application of model-based fault-detection methods in dependence on operation conditions and on data evaluation

		Static behavior		Dynamic behavior	
model type		nonlinear characteristics (steady-state)	nonlinear characteristics (steady-state)	nonlinear dynamic model	nonlinear dynamic model
detection methods		parameter estimation	parity equations	parameter estimation	parity equations
operation conditions	one steady state	–	√	–	–
	different steady states	√	√	–	–
	dynamic excitation	–	–	√	√
	shut-off	–	–	√	√
kind of data evaluation	offline	√	√	√	√
	online (real time)	–	√	√	√

based methods, as e.g. measurement of casing accelerations or structure-borne noise to detect cavitation or an unbalanced rotor as described in Section 6.1.5.

The different methods have different properties with regard to the *kind of data evaluation*, Table 6.7. The parity equation can primarily be used online in real time and give therefore, as signal model-based methods, an immediate information after a fault occurrence. Parameter estimation of the steady-state characteristics can only be applied after gathering all data (batch processing). However, parameter estimation for dynamic operation can be realized online in real time with recursive algorithms.

A feasible application for the fault detection of pumps is to apply parity equations online and in realtime, which does not require a higher computational expense. This can be supplemented by a vibration sensor at the pump casing, but means an additional sensor. If the residuals indicate larger deviations a special dynamic excitation signal (e.g. APRBS) of the speed can be applied for a short time to gain better information on the fault type by dynamic parameter estimation.

6.2 Reciprocating pumps

Reciprocating pumps are used for pumping of various faults, especially within the chemical and pharmaceutical industries. They are also called oscillating positive displacement pumps and are primarily applied for high pressure differences and small to medium flow rates. The pressure can go up to 3000 bar and the volume flow ranges from 0.1 ml/h to 1000 m^3/h. The rated power can reach 1 MW. The piston or diaphragm motion principle also allows one to meter the flow precisely. If oscillating diaphragms are used, aggressive and toxic media can be pumped and metered. As

these pumps frequently have central functions in plants their availability is important, [6.3], [6.27].

6.2.1 Structure of a diaphragm pump

Figure 6.19 shows a scheme of a reciprocating pump. The speed-controlled AC motor is connected to a gear with a crankshaft to generate an oscillating linear motion for a piston in the pump head. This piston moves oil forth and back and therefore generates a displacement of the diaphragm. During the suction phase the spring-loaded inlet valve opens and the outlet valve closes and during the pressure phase they operate oppositely.

Because of the high load for high pressure these pumps may show up a number of faults with increasing operating time. Examples are leakages of the valves, resulting in back flow and lower efficiency, gas enclosures, cavitation in the suction part or faults in the gear and bearings.

For supervision usually only the pressure after the pressure valve or in the oil section is measured and the current of the electrical motor and limit checking is applied. In the following it is shown how a process model and signal-model-based approach allows one to detect small faults and achieve a deeper fault diagnosis, [6.7].

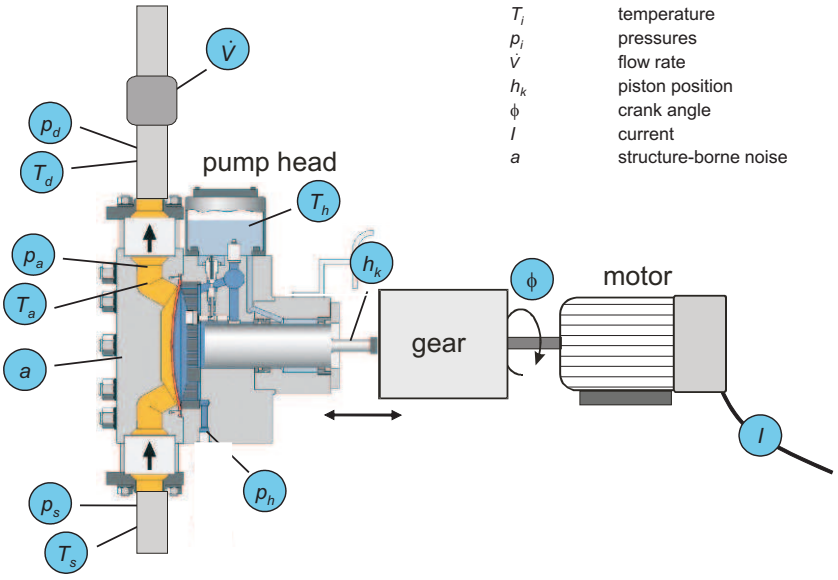


Fig. 6.19. Scheme of a reciprocating diaphragm pump with measured variables

6.2.2 Models of a diaphragm pump

The reciprocating function of the considered pump consists of four phases: pressure build-up, discharge, pressure release, suction. Figure 6.20 shows the pressure behavior dependent on the piston displacement, the so-called indicator diagram. The pressure in dependence on the volume follows for a compressible fluid:

$$dp = -\frac{dV}{V} \tag{6.2.1}$$

where E is the bulk modulus or compressibility module. This module depends on the pressure p, the temperature T and gas contents $\lambda = V_{air}/V_{fluid}$ and is depicted in Figure 6.21. For modeling the pressure one needs a compressibility module E for the pumped fluid and the oil in the hydraulic section. The pressure p_a in the working (pumping) space and the hydraulic space p_h can be assumed as identical:

$$p_a(t) = p_h(t) \tag{6.2.2}$$

The volume behavior follows:

$$V(t) = V_0 - A_p h_p(t) \tag{6.2.3}$$

where V_0 is the total volume of the working and hydraulic space, A_p the piston area and h_p is the piston displacement. Integrating (6.2.1) yields for the build-up phase

$$p_a(t) - p_0 = \int_0^t \frac{E'(p_a, \lambda)}{V_0 - A_p h_p(\tau)} A_p v_p(\tau) d\tau \tag{6.2.4}$$

p_0 is the initial pressure for $t = 0$ and $v_p(t)$ is the piston speed. For E' an analytical expression can be used, [6.7].

Corresponding models ban be given for the discharge stroke taking into account pressure drops at the valves and the pipes and the adiabatic behavior of the pulsation damper, the pressure release and suction stroke. Hence, four models result, which describe the normal behavior of the pump pressure, Figure 6.22. The unknown parameters for all phases are estimated with hill-climbing methods using measured variables of the considered pump. Then the resulting models show a relative good agreement with measurements, [6.7].

6.2.3 Fault detection and fault diagnosis of the hydraulic pump

a) Pressure signal

Some faults change the pressure behavior $p_a(h_p)$. For example, gas enclosure lead to a retarded pressure increase and a leakage in the pressure valve to an advanced pressure increase. Therefore the differences between different crankshaft angles for the $p_a(\phi)$ signal for the normal and the observed behavior lead to four symptoms and the integration of the output pressure and measured pressure to three other symptoms,

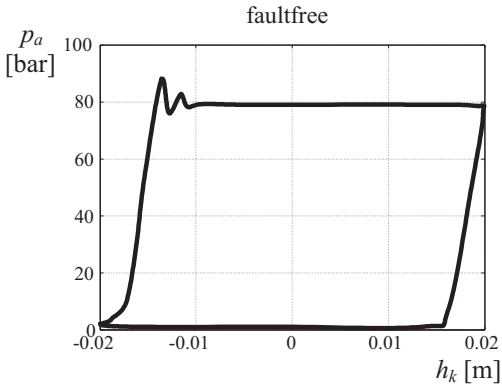


Fig. 6.20. Pressure behavior for one cycle (indicator diagram)

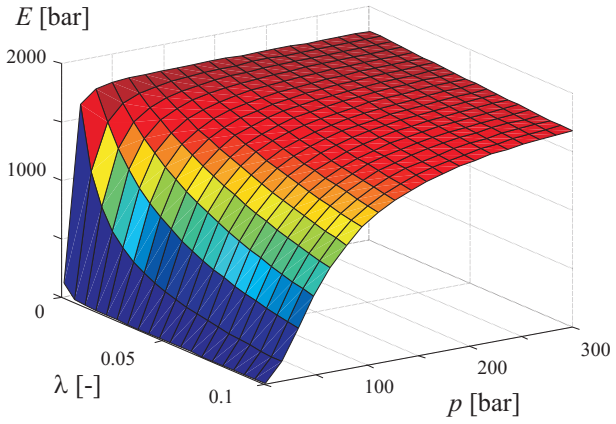


Fig. 6.21. Compressibility module in dependence on pressure and gas contents for hydraulic oil

Figure 6.23. The last symptoms integrate the pressure difference over the crank angle, resulting in deviation areas. This means that the difference between the pump model and the real behavior are generated, in the sense of parity equations, but special features are calculated which reflect faults.

For experimental testing of the fault-detection method a pump test bench was installed, Figure 6.24, where the variables indicated in Figure 6.19 could be measured.

b) Acceleration signal

Additional information can be gained by an accelerometer at the pump head, see Figure 6.19. Especially for leaks at the pressure and the suction valve structure-borne noise can be observed in special time windows. Therefore the variance of the sampled accelerometer signal $a(t)$ is calculated:

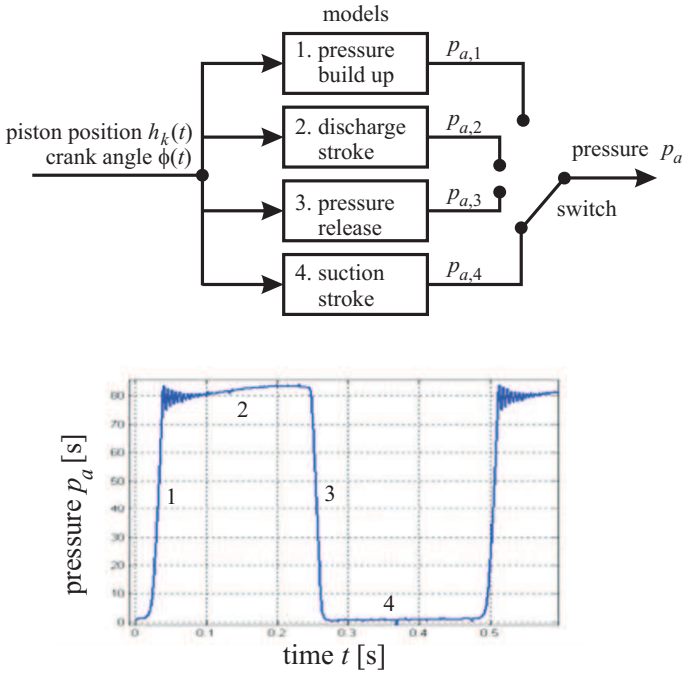


Fig. 6.22. Models for the outlet pressure

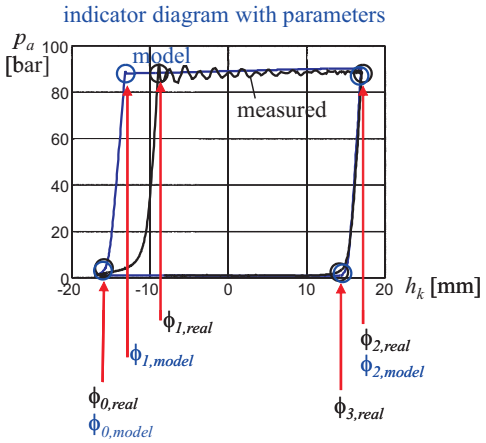
$$\sigma_a^2 = \frac{1}{N-1}\sum_{k=1}^{N}[a(k)-\bar{a}]^2 \tag{6.2.5}$$

with sampling time $T_0 = t/k = 1$ ms. Then two symptoms result for the discharge and suction phase:

$$\begin{aligned} S_{10} &= \sigma^2_{a,\text{disch}}(a) - \sigma^2_{\text{ref,disch}}(a) \\ S_{11} &= \sigma^2_{a,\text{suc}}(a) - \sigma^2_{\text{ref,suc}}(a) \end{aligned} \tag{6.2.6}$$

The increasing noise results from the fluid pressed through the leakages of the valves under high pressure differences leading to strong flow noise and to cavitation if pressure falls under vapor pressure, see Figure 6.25. The evaluation of the accelerometer signal is performed only for time windows of the discharge and suction phase, as shown in Figures 6.25 and 6.26.

Figure 6.27 depicts the overall scheme for the fault detection with 11 symptoms. The resulting signs and sizes of the symptoms for different faults are presented in Table 6.8. They show different patterns for the investigated hydraulic faults and therefore allow a clear distinction and indication of the fault sizes. The final fault diagnosis was implemented with fuzzy if-then rules, [6.7]. The smallest faults which could be detected correspond to a decrease of the volumetric efficiency of the pump of about 2%. Also some multiple faults can be detected.



1. Crank angle symptoms

$$S_1 = \Delta\phi_0 = \phi_{0,model} - \phi_{0,real}$$
$$S_2 = \Delta\phi_1 = \phi_{1,model} - \phi_{1,real}$$
$$S_3 = \Delta\phi_2 = \phi_{2,model} - \phi_{2,real}$$
$$S_4 = \Delta\phi_3 = \phi_{3,model} - \phi_{3,real}$$

2. Pressure symptoms

$$S_5 = \frac{1}{\phi_{1,real} - \phi_{0,real}} \cdot \int_{\phi_{0,real}}^{\phi_{1,real}} \left(p_{a,model} - p_{a,real}\right) d\phi$$
$$S_6 = \frac{1}{\phi_{3,real} - \phi_{2,real}} \cdot \int_{\phi_{2,real}}^{\phi_{3,real}} \left(p_{a,model} - p_{a,real}\right) d\phi$$
$$S_7 = \frac{1}{2\pi - \phi_{3,real}} \cdot \int_{\phi_{3,real}}^{2\pi} \left(p_{a,model} - p_{a,real}\right) d\phi$$

Fig. 6.23. Symptoms generated from $p(h_p)$-behavior

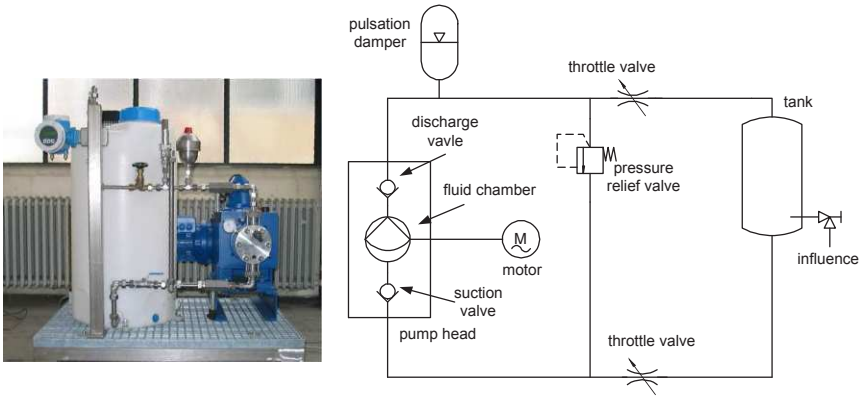


Fig. 6.24. Test bench for a reciprocating pump $p_{max} = 100$ bar, $\dot{V}_{max} = 342$ l/h, max speed $n = 200$ rpm, 4 kW AC motor, manual stroke setting

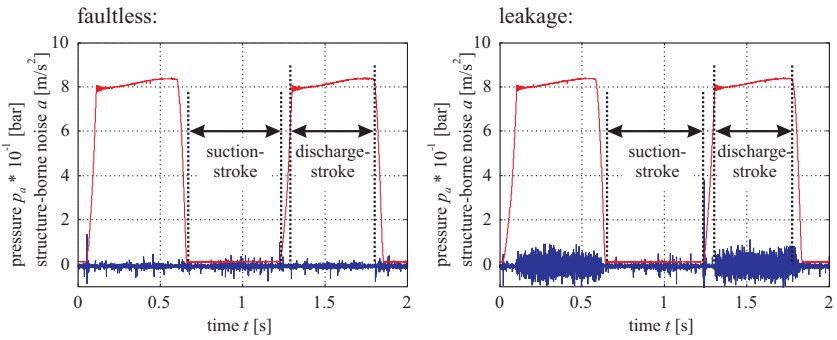


Fig. 6.25. Structure-borne noise for a leakage at the suction valve

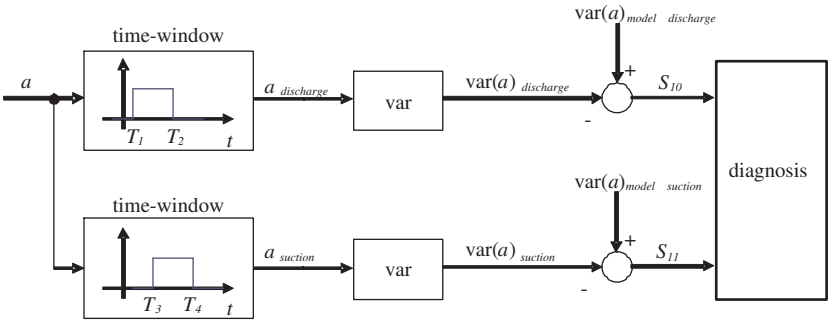


Fig. 6.26. Generation of time-windowed variances of the structure-borne noise

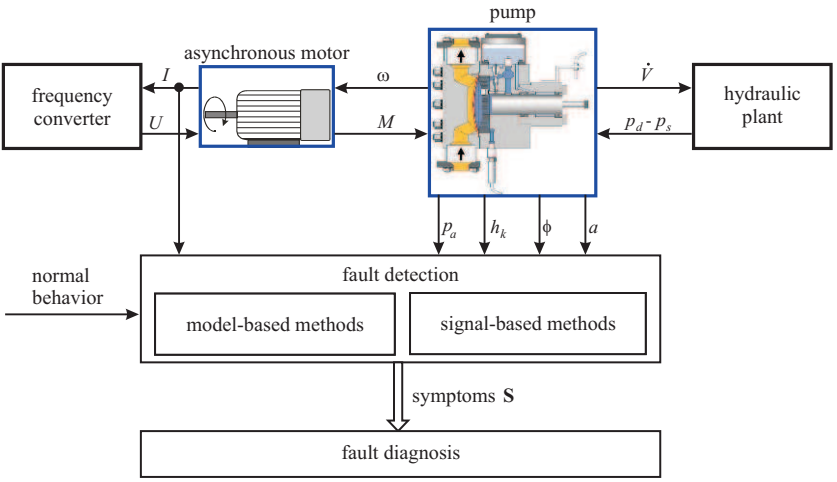


Fig. 6.27. Overall scheme for fault detection and diagnosis of the reciprocating pump

Table 6.8. Fault-symptom table for different faults and fault sizes

		Crank angle symptoms				Pressure symptoms			Pressure increase symptoms		Structure-borne noise symptoms	
Gas in fluid	small	-	-	0	0	+	0	0	0	0	0	0
	medium	–	–	0	0	++	0	0	0	0	0	0
	large	—	—	0	0	+++	0	0	0	0	0	0
Throttle & cavitation	small	-	-	0	0	+	0	+	0	0	0	0
	medium	–	–	0	0	++	0	++	0	0	0	0
	large	—	—	0	0	+++	0	++	0	0	0	0
Leakage pressure valve	small	0	0	0	0	0	0	+	0	0	0	++
	medium	0	+	0	-	-	-	0	-	+	0	+++
	large	–	++	0	–	–	–	0	++	++	0	+++
Leakage suction valve	small	0	0	0	0	0	0	+	0	0	++	0
	medium	0	-	0	+	+	+	0	+	-	+++	0
	large	0	–	++	++	++	++	0	++	++	+++	0
		pressure signal									structure-borne noise	

6.2.4 Fault detection of the pump drive

To detect faults in the drive chain of the pump the following measured signals can be used: the effective phase current I_S of the AC motor, the angle velocity ω_{rot} of the AC motor, the piston displacement h_p and the pump pressure p_a. The torque balance at the motor shaft is

$$J\dot{\omega}_{rot}(t) = M_{el}(t) - M_{pump}(t) - M_f(t) \tag{6.2.7}$$

where J is the ratio of inertia, M_{el} the motor torque, M_{pump} the pump torque and M_R the friction torque of the bearings, gear and crankshaft. The torque M_{el} of the electrical motor can be calculated from simplified equations and the measured effective value of the phase current $I_s(t)$, see, e.g. [6.4], [6.28], [6.14]. The friction torque is assumed as

$$M_f(t) = M_c(t) + c_v\omega_{rot}(t) \tag{6.2.8}$$

consisting of a Coulomb-friction coefficient M_c and viscous-friction coefficient c_v.

The load torque resulting from the pump is

$$M_{pump}(t) = M_{hyd}(t) + M_{fric}(t) + M_{mass}(t) \tag{6.2.9}$$

with

$$
\begin{aligned}
M_{hyd} &= \frac{1}{\omega'_{rot}} A_p \dot{h}_p (p_a - p_{atm}) \\
M_{fric} &= \frac{1}{\omega'_{rot}} (F_p + c_{p,v} \dot{h}_k) \dot{h}_k \\
M_{mass} &= \frac{1}{\omega'_{rot}} m_k \ddot{h}_k \dot{h}_k \\
\omega'_{rot} &= \omega_{rot} / i \quad (i = \text{gear ratio})
\end{aligned}
$$

Mostly $M_{mass} \ll M_{hyd}$ and can therefore be neglected. Based on these models a residual for the drive chain friction can be calculated:

$$
\begin{aligned}
r_f(t) &= M_f(t) - M_{f,ref}(t) \\
&= M_f(t) - M_c(t) - c_v \omega_{rot}(t)
\end{aligned}
\tag{6.2.10}
$$

where the observed friction torque follows from (6.2.9), Figure 6.28. If this residual passes a certain threshold an increased friction is observed, e.g. generated by faults in the bearings, gear or crankshaft. Figure 6.29 indicates how the residual changes after increasing the friction torque. Experiments have shown that faults which change the torque by about 10% can be detected with the presented model-based approach, [6.7].

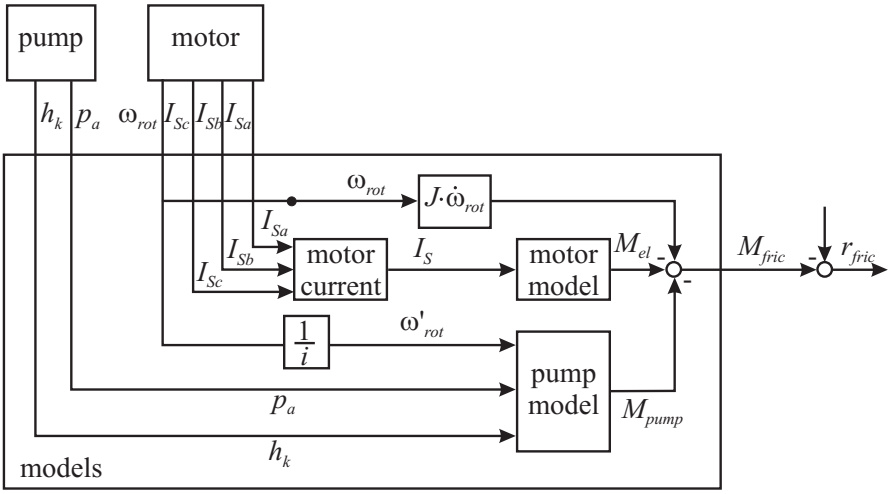


Fig. 6.28. Signal flow for the model-based fault detection of the drive train

6.2.5 Conclusions

An overall scheme of the model-based fault detection and diagnosis of the investigated reciprocating pump in depicted in Figure 6.27. It could be demonstrated with

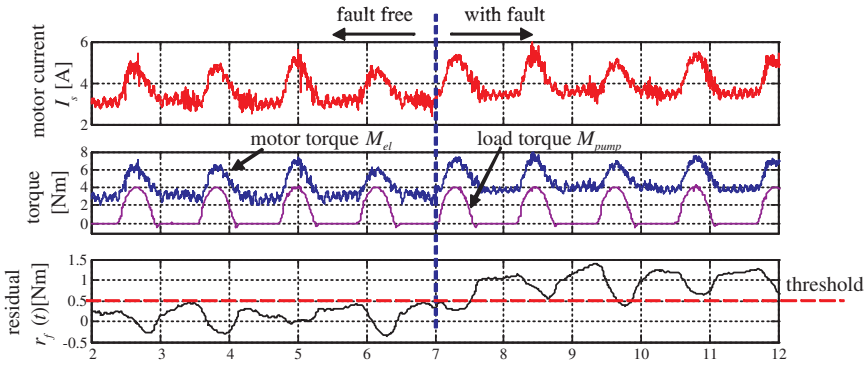


Fig. 6.29. Measured motor current I_s, calculated motor torque M_{el} and pump torque M_{pump} and friction residual r_f for an increase of friction at $t = 7\,\mathrm{s}$

the experiments at the test bench that special features calculated from the pump pressure behavior and residuals of the structure-borne noise allow a detailed diagnosis of small faults within the hydraulic part. In addition by measuring the current and speed of the electrical motor faults in the mechanical part leading to an increased friction can be detected using the "drive-as-sensor principle". Thus, a combination of process-model and signal-model methods could be successfully applied to this pump type.

7

Leak detection of pipelines

The leak detection of pipelines is a basic task for any pipeline operation. Pipelines consist usually of different sections with a length of 30–100 km and a pump or a compressor at its inlet and a tank or storage at the outlet, Figure 7.1. The available measurements are mostly pressure p_0 and p_l, mass flow $\dot{m}_0$ and $\dot{m}_l$ and temperature T_0 and T_l at inlet and outlet of the pipeline or a pipeline section. In some cases the sections are separated by a sliding valve with additional pressure measurements. The measured signals are transmitted to a control station by cables, optic fibres or wireless communication, sometimes as redundant lines. With regard to the leak detection and leak localization several cases have to be taken into account:

(i) medium: liquid – gas – multiple phases
(ii) operation: standstill – stationary – non-stationary
(iii) size of leak: small – medium – large
(iv) timely development of leak: abrupt (cracking of welding seam), slowly (hole corrosion) or already existing
(v) leak monitoring: continuously – in time intervals – on request.

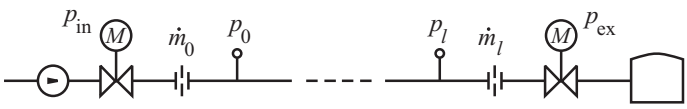


Fig. 7.1. Usual instrumentation of a pipeline or pipeline section. p pressure, $\dot{m}$ mass flow

7.1 State of the art in pipeline supervision

The following ways are possible to detect leaks during operation:

- *Pressure probe*: A pressurized pipeline section is closed by sliding valves and the time-dependent pressure development is observed. This is a very sensitive

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_7, © Springer-Verlag Berlin Heidelberg 2011

method, but it requires a standstill of pumping and leak localization is not possible.

- *Mass balance*: Taking the difference of input and output mass flow

$$\dot{m}_L(t) = \dot{m}_0(t) - \dot{m}_l(t) \tag{7.1.1}$$

allows one to detect leaks in liquid pipelines of about $\dot{m}_L(t) \geq 2\%$ and in gas pipelines $\geq 10\%$. Due to inherent dynamics, noise effects and measurement defects no smaller leaks can usually be detected and their localization is not possible.

- *Pressure wave detection*: Using highly sensitive and fast pressure sensors shock waves can be detected after larger, abrupt leaks in liquid pipelines. Based on the known speed of sound of about 800–1500 m/s, localization is possible with an accuracy of about 100 m, [7.9].
- *Ultrasonic noise*: The sound of a leak is especially transmitted through the pipe wall and can be measured by microphones at various locations along the pipeline. Time shifts of the signals then allow one to estimate the leak location approximately.
- *Liquid-sensitive cables*: If the pipeline is equipped with liquid-sensitive cables at the bottom, the resistance of the cables is changed in the case of a leak, what can be used for leak detection and localization by analyzing a network consisting of different serially connected resistances.
- *Process-model-based methods*: The stationary and dynamic behavior of the pressures and mass flow rates are described by a mathematical pipeline model and are compared with measured variables. Thus, residuals are generated and parameters are estimated to allow leak detection and localization as well for liquids as for gases. This will be considered in the sequel. These methods were first published by [7.8], [7.7] and [7.3].

7.2 Models of pipelines

A simplified scheme of a pipeline according to Figure 7.2 is considered. The physical data are:

z	length coordinate
l	length of the pipeline
d_F	inner diameter of the pipe
$A_F = \pi d_F^2/4$	cross-sectional area of the pipe
$H(z)$	height profile of the pipeline
$p(z,t)$	fluid pressure
$\rho(z,t)$	fluid density
$T(z,t)$	fluid absolute temperature
$w(z,t)$	fluid velocity
m	fluid mass
$\dot{m}(z,t)$	fluid mass flow rate

$\dot{m}_L$	leak mass flow rate
R	gas constant
$c_F = \sqrt{p/\rho}$	speed of sound
λ	friction coefficient
ν_f	fluid viscosity
$\dot{V}$	volume flow rate

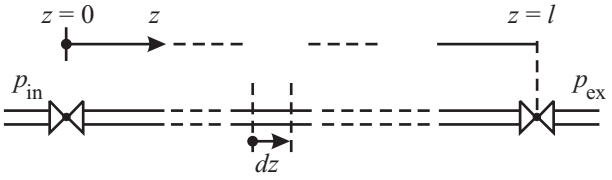


Fig. 7.2. Simplified scheme of a pipeline (H = const.)

For a pipe element of length dz the *mass balance equation* becomes

$$\frac{\partial m}{\partial t} = A_F \rho w - A_F \left(w + \frac{\partial w}{\partial z} dz \right) \left(\rho + \frac{\partial \rho}{\partial z} dz \right) \tag{7.2.1}$$

and with

$$\dot{m} = A_F \rho w \tag{7.2.2}$$

and neglect of small terms

$$\frac{\partial}{\partial z}(\rho w) + \frac{\partial \rho}{\partial t} = 0 \tag{7.2.3}$$

The *momentum balance* is

$$\begin{aligned} \frac{\partial}{\partial t}(A_F \rho w dz) =& A_F \left(p + \frac{\rho w^2}{2} \right) - A_F \left(p + \frac{\partial p}{\partial z} dz + \frac{\rho w^2}{2} \right. \\ &\left. + \frac{\partial}{\partial z} \frac{(\rho w^2)}{2} dz \right) - A_F F - A_F Y \end{aligned} \tag{7.2.4}$$

or

$$\frac{\partial}{\partial t}(\rho w) + \frac{\partial}{\partial z} \left(p + \frac{\rho w^2}{2} \right) = -F - Y \tag{7.2.5}$$

The friction force related to A_F is

$$F = \frac{\partial p_F}{\partial z} = \lambda \frac{\rho}{2 d_F} w |w| \tag{7.2.6}$$

where λ is the friction coefficient. It depends on the Reynolds number

$$Re = \frac{d_F w}{\nu_f} \tag{7.2.7}$$

Usually laminar flow can be assumed for $Re < 2320$ with

$$\lambda = \frac{64}{Re} \tag{7.2.8}$$

and turbulent flow for $Re > 2320$ with

$$\lambda = \frac{0.3164}{Re^{0.25}} \tag{7.2.9}$$

The static pressure term becomes

$$Y = \rho g \frac{dH}{dz} = \rho g \sin\alpha \tag{7.2.10}$$

if α is the angle of ascent.

If an isothermic flow with temperature T_0 can be assumed the gas state equation is

$$p\frac{1}{\rho} = Z(p, T_0)RT_0 = c_F^2(p) \tag{7.2.11}$$

where $c_F(p)$ is the isothermic speed of sound. Then after the introduction of the variables $\dot{m}(z,t)$ and $p(z,t)$ the two balance equations become

$$A_F \frac{\partial}{\partial t}\left(\frac{1}{c_F^2(p)}p\right) + \frac{\partial \dot{m}}{\partial z} = 0 \tag{7.2.12}$$

$$\begin{aligned} &\frac{1}{A_F}\frac{\partial \dot{m}}{\partial t} + \frac{\partial}{\partial z}\left(p + \frac{\dot{m}^2 c_F^2(p)}{2A_F^2 p}\right) = \frac{1}{A_F}\frac{\partial \dot{m}}{\partial t} \\ &+ \left(1 - \frac{\dot{m}^2 c_F^2}{2A_F^2 p^2}\right)\frac{\partial p}{\partial z} + \frac{\dot{m} c_F^2 \partial \dot{m}}{A_F^2 p \partial z} = -F - Y \end{aligned} \tag{7.2.13}$$

$$F = \frac{\lambda}{2d_F}\frac{c_F^2(p)}{A_F^2 p}\dot{m}|\dot{m}|$$

$$Y = \frac{gp}{c_F^2(p)}\frac{dH}{dz}$$

This pipeline model can now be simplified by assuming:

(i) the isothermic speed of sound is constant within a pipeline section $j : c_F(p) = c_{Fj}$

(ii) the fluid flow velocity w_F is small in comparison to the speed of sound c_F, so that

$$w_F^2 \ll c_F^2$$

$$\text{i.e. } \frac{\dot{m}^2 c_F^2}{A_F^2 p^2} = \frac{w_F^2}{c_F^2} \approx 0$$

(iii) for slow dynamic changes and long pipelines the term

$$\frac{\dot{m}c_F^2}{A_F^2 p}\frac{\partial \dot{m}}{\partial z} \approx 0$$

can be neglected.

Then, the simplified "long-pipeline model" results in the assumption $\partial H/\partial z = 0$:

$$k_1 \frac{\partial p}{\partial t} + \frac{\partial \dot{m}}{\partial z} = 0 \tag{7.2.14}$$

$$k_2 \frac{\partial \dot{m}}{\partial t} + \frac{\partial p}{\partial z} = -k_3 \dot{m}\frac{|\dot{m}|}{p} \tag{7.2.15}$$

which is a hyperbolic partial differential equation system, with the coefficients

$$\begin{aligned} k_1 &= \frac{A_F}{c_{Fj}^2} \\ k_2 &= \frac{1}{A_F} \\ k_3 &= \frac{\lambda}{2d_F}\frac{c_{Fj}^2}{A_F^2} \end{aligned} \tag{7.2.16}$$

For the solution of the partial differential equation system the pipeline is subdivided (discretized) into sections j, Figure 7.3, so that

$$\begin{aligned} \frac{\partial p_j}{\partial t} &= g_{1j}\left(\dot{m}_{j+1} - \dot{m}_{j-1}\right) \quad j = 1, 3, \ldots, l-1 \\ \frac{\partial \dot{m}_j}{\partial t} &= g_2\left(p_{j+1} - p_{j-1}\right) + g_{3(j-1)}\dot{m}_j|\dot{m}_j| \quad j = 2, 4, \ldots, l-2 \\ \frac{\partial \dot{m}_0}{\partial t} &= g_{20}\left(p_1 - p_0\right) + g_{30}\dot{m}_0|\dot{m}_0| \\ \frac{\partial \dot{m}_1}{\partial t} &= g_{21}\left(p_l - p_{l-1}\right) + g_{31}\dot{m}_l|\dot{m}_l| \quad (l = \text{even}) \end{aligned} \tag{7.2.17}$$

with

$$\begin{aligned} g_{1j} &= -\tfrac{1}{k_1 \Delta z} & g_{20} &= -\tfrac{2}{k_2 \Delta z} \\ g_2 &= -\tfrac{1}{k_2 \Delta z} & g_{21} &= -\tfrac{2}{k_2 \Delta z} \\ g_{3(j-1)} &= -\tfrac{k_{3j}}{p_{j-1}} & g_{30} &= -\tfrac{k_3}{p_0} \\ \Delta z &= \tfrac{2L}{l} & g_{31} &= -\tfrac{k_3}{p_{l-1}} \end{aligned} \tag{7.2.18}$$

The boundary conditions are given by the valve equations

$$\begin{aligned} \dot{m}_0 &= k_{v0}\sqrt{\left(\frac{\rho_{00}}{p_{00}}\rho_0\right)}\sqrt{[(p_{in} - p_0)]} = c_{v0}\sqrt{(p_{in} - p_0)} \\ \dot{m}_l &= k_{vl}\sqrt{\left(\frac{\rho_{00}}{p_{00}}\rho_l\right)}\sqrt{[(p_l - p_{ex})]} = c_{vl}\sqrt{(p_l - p_{ex})} \end{aligned} \tag{7.2.19}$$

with ρ_{00} and p_{00} as reference values of water and k_v as the k_v-values of a valve. (7.2.17) can now be expressed as a nonlinear state-variable representation

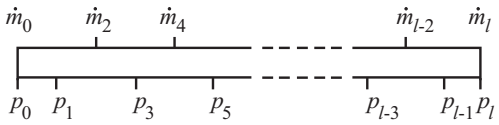


Fig. 7.3. Subdivision of the pipeline into sections

$$\begin{aligned} \dot{x}(t) &= A(x)x(t) + Bu_p(t) \\ y(t) &= Cx(t) \end{aligned} \tag{7.2.20}$$

with

$$x^T = \left[\dot{m}_0 \dot{m}_2 \ldots \dot{m}_l \vdots p_1 p_3 \ldots p_{l-1}\right]$$
$$u_p^T = [p_0 p_l] \quad y^T = [\dot{m}_0 m_l]$$

$$A = \begin{bmatrix} g_{30}|\dot{m}_0| & 0 & \ldots & 0 & g_{20} & 0 \ldots 0 \\ 0 & g_{31}|\dot{m}_2| & \ldots & 0 & -g_2 & g_2 \; 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \; \vdots \\ -g_{11} & g_{11} & \ldots & 0 & 0 & \ldots 0 \\ 0 & -g_{13} & g_{13} & 0 & 0 & \ldots 0 \\ \vdots & \vdots & \vdots & \vdots & & \\ 0 & \ldots & -g_{1(l-1)} & g_{1(l-1)} & 0 & \ldots 0 \end{bmatrix} \tag{7.2.21}$$

$$B = \begin{bmatrix} -g_{20} & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & g_{21} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}$$

The considered equations can also be used for *compressible liquids*. Then in (7.2.13) the coefficient $\dot{m}|\dot{m}|$ becomes

$$\frac{k_3}{p} = \frac{\lambda}{2d_F A_F^2} \frac{c_{Fj}^2}{p_j} = \frac{\lambda}{2d_F A_F^2} \rho_j \tag{7.2.22}$$

and therefore $g_{3(j-1)}$ becomes a constant for each section j.

If a small *leak flow* $d\dot{m}_{L\xi}$ occurs at section $j = \xi$, see Figure 7.4, this can be modeled by introducing it into the mass balance equation (7.2.3):

$$\frac{\partial}{\partial z}(\rho w)_\xi + \frac{\partial \rho_\xi}{\partial t} + \frac{1}{A_F}\frac{\partial \dot{m}_{L\xi}}{\partial z} = 0 \tag{7.2.23}$$

Then (7.2.14) changes to

$$k_1 \frac{\partial p_\xi}{\partial t} + \frac{\partial \dot{m}}{\partial z} + \frac{\partial \dot{m}_{L\xi}}{\partial z} = 0 \tag{7.2.24}$$

and (7.2.17) to

$$\frac{\partial p_\xi}{\partial t} = g_{1\xi}\left(\dot{m}_{\xi+1} - \dot{m}_{\xi-1}\right) + g_{1\xi}\dot{m}_{L\xi} \tag{7.2.25}$$

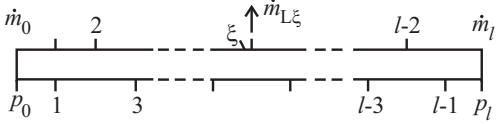


Fig. 7.4. Leak flow $\dot{m}_{L\xi}$ at section ξ

7.3 Model-based leak detection

Based on the mathematical pipeline model, leak-detection methods are considered for liquids and gases, non-stationary or slow dynamic operation with small changes of the variables and small leaks which may appear abruptly or slowly, and for continuous monitoring.

For small changes and the mass flow rate $\dot{m}$ in the positive z-direction, the momentum balance equation of (7.2.17) can be linearized:

$$\frac{\partial \dot{m}_j}{\partial t} = g_2(\Delta p_{j+1} - \Delta p_{j-1}) + 2g'_{3(j-1)}\Delta \dot{m}_j \tag{7.3.1}$$

where all coefficients are taken for the steady-state values $\bar{p}_j$ and $\bar{\dot{m}}_j$. Further the linearized valve equations (7.2.19) are introduced in the form

$$\left.\begin{aligned} \Delta p_0 &= c'_{v0}\Delta \dot{m}_0 + \Delta p_{in} \\ \Delta p_l &= c'_{vl}\Delta \dot{m}_l + \Delta p_{ex} \end{aligned}\right\} \tag{7.3.2}$$

Then a linear state representation

$$\left.\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}\right\} \tag{7.3.3}$$

results with

$$\left.\begin{aligned} x^T(t) &= \left[\Delta\dot{m}_0 \Delta\dot{m}_2 \ldots \Delta\dot{m}_l \vdots \Delta p_1 \Delta p_3 \ldots \Delta p_{l-1}\right] \\ u^T &= [\Delta p_{in} \Delta p_{ex}] \\ y^T &= [\Delta\dot{m}_0 \Delta\dot{m}_l] \end{aligned}\right\} \tag{7.3.4}$$

However, for most of the *gas pipelines* the nonlinear (7.2.20) must be used. This results from their big storage capacities and the time-dependent consumption. Therefore, gas pipelines rarely come to a steady state.

7.3.1 Leak detection with state observers

It is assumed that a small leak flow $\delta\dot{m}_L$ occurs at section $j = \xi$. The effect of the leak can be modeled by introducing this leak mass flow rate into the mass balance of this section, see (7.2.23) leading to (7.2.25). This changes the linearized state equation (7.3.3) to

$$\dot{x}(t) = Ax(t) + Lv(t) + Bu(t) \tag{7.3.5}$$

with the leak flow vector

$$v^T(t) = \left[0 \quad 0 \ldots \dot{m}_{l\xi} \ldots 0 \quad 0 \ldots 0\right] \tag{7.3.6}$$

and the leak influence matrix

$$L = \begin{bmatrix} 0 & \ldots & 0 & 0 \ldots 0 \\ \vdots & & \vdots & \vdots \quad \vdots \\ 0 & \ldots & 0 & 0 \ldots 0 \\ 0 & g_{11} \ldots & 0 & 0 \ldots 0 \\ \vdots & g_{1\xi} & \vdots & \vdots \quad \vdots \\ 0 & \ldots & g_{1(m-1)} & 0 \ldots 0 \end{bmatrix} \tag{7.3.7}$$

Hence, a leak flow appears as a disturbance or unknown input variable of the state variables $x(t)$, Figure 7.5.

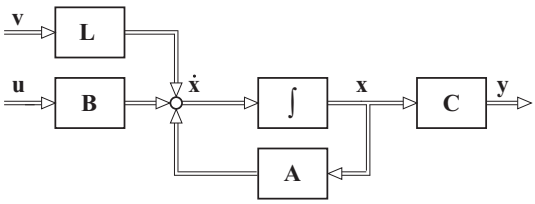


Fig. 7.5. A leak appears as disturbance $v(t)$

The leak monitoring task now consists in the detection of an appearing leakage, its localization $\hat{z}_L$ along the pipeline and the estimation of its size $\dot{m}_L$. In most cases

only measurements $\dot{m}_0(t)$, $p_0(t)$ and $\dot{m}_l(t)$, $p_l(t)$ at the inlet and the exit of the pipeline are available.

The results of simulations for a gasoline and an ethylene-gas pipeline assuming different locations of a suddenly appearing leak of 5% of the mass flow rate are shown in [7.11]. If the leak location is approximately in the middle of the pipeline the flows $\dot{m}_0(t)$ and $\dot{m}_l(t)$ change with about the same settling times and reach their new steady states after about 4 min for the gasoline pipeline and 2 h for the gas pipeline. If the leak is closer to one end, the time responses and the magnitudes of the flow changes become rather different. Now various approaches for leak monitoring are shown.

a) State observer as fault-sensitive filter

For the detection of the leak a state observer

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + H\,[y(t) - C\hat{x}(t)] \tag{7.3.8}$$

can be designed to reconstruct the states and to calculate the residuals

$$\tilde{y}(t) = y(t) - C\hat{x}(t) \tag{7.3.9}$$

thereby it is assumed that all parameters of the pipeline system are known.

In the first adjustment phase, the filter gain H can be large so that a fast adjustment occurs. For leak detection the gain H is lowered. Then if a leak $\dot{m}_{L\xi}$ occurs suddenly the residuals $\tilde{y}$ show a deviation $+\Delta\dot{m}_0(t)$ and $-\Delta\dot{m}_l(t)$, i.e. the residuals change in predetermined directions that can be used for the fault detection.

One drawback of this method is that the filter tries to adapt to the leak-disturbed process after a while so that the information on the leak vanishes with time.

b) State observer with fault model

Another method consists of reconstructing the leak flow vector $\nu(t)$ by modeling the fault influence in the filter, see Figure 7.6. The influence of a suddenly (stepwise) appearing leak is modeled by

$$\dot{\nu} = \nu(t) \tag{7.3.10}$$

with initial value $\nu(0) = 1$ and $\dot{\nu}(t) = 0$ for $t > 0$ so that the observer is able to reconstruct a remaining leak vector

$$\hat{\nu} = H_\nu \int_{t_1}^{t} \tilde{y}(t')dt' \tag{7.3.11}$$

If the observer converges in the right way the estimated leak vector contains the size as well as the section number ξ where the leak occurred. To extract this information under noisy conditions a bank of filters could be used, assuming different locations.

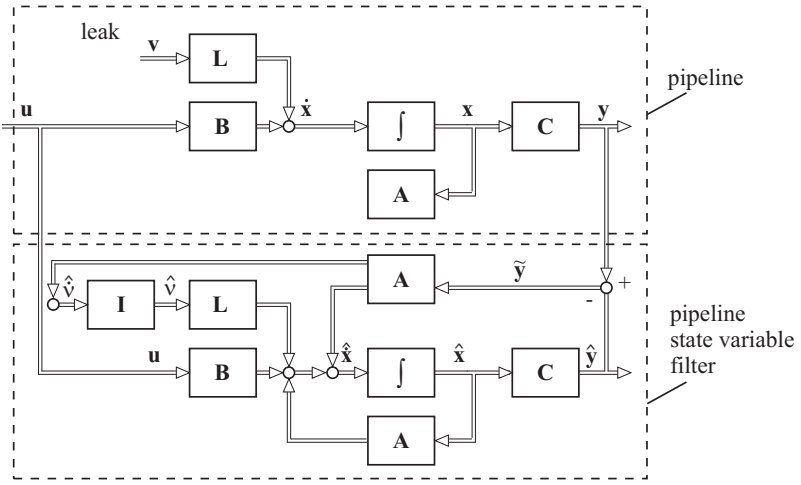


Fig. 7.6. A leak appears as disturbance $v(t)$

A simulation with multiple model hypothesis probability testing for an oil pipeline of 30 km length, using parallel Kalman filters, is described by [7.6]. A leak of 1% of the total flow was detected after about 160 s.

Another reference on a fault-sensitive state-variable detector is [7.5]. Here, a simulation study on the detection of diverted or stolen nuclear material (corresponding to a leak) in a plutonium concentrator by using statistical decision methods for the residuals of an extended Kalman filter is shown. Relatively large computation time and storage was required.

In order to estimate the leak location for a pipeline of 100 km length with an accuracy of about ± 1 km at least 50 sections have to be modeled so that a high system order results which may cause computational problems. Additionally, several process parameters are not known precisely enough, i.e. the friction coefficients, and also several temperature effects, so that the models have to be updated by parameter estimation methods, [7.4].

Therefore, a simpler method was developed for liquid pipelines, taking into account several practical requirements and special cases of pipeline operation.

7.3.2 Leak detection with mass balance and correlation analysis for liquid pipelines

It is assumed that a liquid pipeline operates in a stationary steady state and that only the mass flow rates at the inlet $\dot{m}_0(k)$ and the exit $\dot{m}_l(k)$ can be measured, where k is the discrete time.

The simplest method of leakage monitoring is then by stating a static balance equation

$$\dot{m}_L(k) = \dot{m}_0(k) - \dot{m}_l(k) \tag{7.3.12}$$

and by triggering an alarm if the leak flow $\dot{m}_L$ exceeds a certain limit. However, this pure balancing method is not suitable for the detection of smaller leaks, because of the noise signals, drifting measurements and the dynamic changes of both flows.

An improvement of this method is obtained by determining the *low-frequency components* of the flows by discrete-time low-pass filtering:

$$\dot{m}_j^*(k) = \kappa_m \dot{m}_j^*(k-1) + (1-\kappa_m)\dot{m}_j^*(k) \quad (j = 0, l) \tag{7.3.13}$$

$\dot{m}_0^*(k)$ and $\dot{m}_l^*(k)$ are then reference values, which may be different because of calibration errors of the sensors. They also change slowly due to temperature and viscosity changes in the case of constant pumping. A leak is then obtained by

$$\begin{aligned} \Delta\dot{m}_0(k) &= \dot{m}_0(k) - \dot{m}_0^*(k) \\ \Delta\dot{m}_l(k) &= \dot{m}_l(k) - \dot{m}_l^*(k) \\ \dot{m}'_L(k) &= \Delta\dot{m}_0(k) - \Delta\dot{m}_l(k) \end{aligned} \tag{7.3.14}$$

or by further low-pass filtering:

$$\dot{m}''_L(k) = \kappa_L \dot{m}'_L(k-1) - (1-\kappa_L)\dot{m}'_L(k) \tag{7.3.15}$$

where $\kappa_L < \kappa_m$ is required in order to be able to detect suddenly appearing leaks. If $\dot{m}''_L(k)$ exceeds a certain threshold

$$\dot{m}''_L(k) > \dot{m}_{Lth}$$

a *leak alarm* is given.

By the described low-pass filtering, noise effects and slow drift effects can be partly eliminated, but changes of the flows according to the inherent fluid dynamics cause the adjustment of relatively large thresholds $\dot{m}_{Lth}$ in order to avoid too frequent false alarms.

This is one of the reasons to *cross-correlate* the differences:

$$R_{MM}(\tau) = \frac{1}{N}\sum_{k=1}^{N} \Delta\dot{m}_0(k-\tau)\Delta\dot{m}_l(k) \tag{7.3.16}$$

see [7.8], [7.12]. This cross-correlation function reacts sensitively even to small leaks, reduces noise effects and models inherent dynamic relationships between the flow changes. For *further noise reduction* the correlation function is averaged with respect to

$$R_\Sigma = \frac{1}{2P+1}\sum_{\tau=-P}^{P} R_{MM}(\tau) \tag{7.3.17}$$

After a leak has occurred the *fault symptom* consists of changes in the predetermined directions $+\Delta\dot{m}_0$ and $-\Delta\dot{m}_l$, such that the products become negative and R_Σ decreases, see Figure 7.7. An alarm is given, if

$$R_\Sigma < R_{\Sigma\varepsilon} \tag{7.3.18}$$

The cross-correlation function is calculated recursively with forgetting memory:

$$R_{MM}(\tau, k) = \lambda R_{MM}(\tau, k-1) + (1-\lambda)\,[\Delta\dot{m}_0(k-\tau)\Delta\dot{m}_l(k)] \qquad (7.3.19)$$

where $0.9 < \lambda < 1$. Larger values of λ result in improved smoothing and thus reduce the noise, but lead in turn to a delayed alarm.

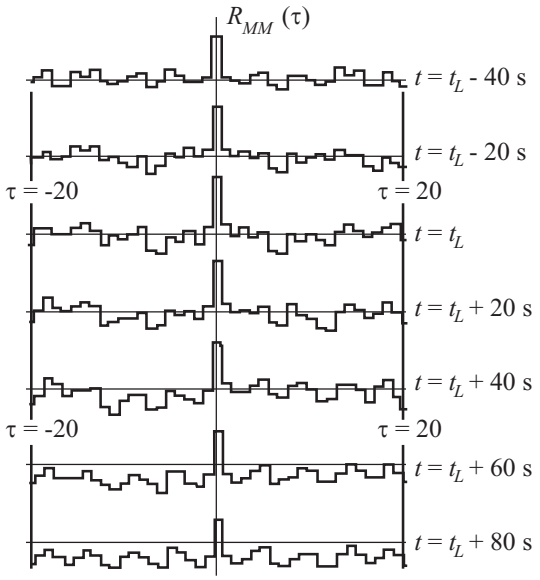


Fig. 7.7. Cross-correlation function $R_{MM}(\tau)$ of the measured inlet and outlet flows after appearance of a leak after t_L. The CCF shifts to negative values because statistically more negative than positive products appear

After the leak has been detected the *leak location* and the size of the *leak mass flow rate* have to be estimated. Assuming stationary operating conditions and measurements of the pressures p_0 and p_l at the inlet and outlet it follows for the pressure gradient before the leak, compare Figure 7.8,

$$\frac{\partial p}{\partial z} = p_z = \frac{p_0 - p_l}{l} = \frac{\dot{m}^2}{p_0 - p_l} \qquad (7.3.20)$$

and after the leak for the upstream gradient

$$p_{zI}^{L} = \frac{\dot{m}_0^2}{p_0 - p_L} \qquad (7.3.21)$$

and downstream

$$p_{zII}^{L} = \frac{\dot{m}_l^2}{p_L - p_l} \qquad (7.3.22)$$

where the (unknown) pressure at the leak location z_L is

$$p_L = p_0 - p_{zI}^L z_L = p_l + p_{zII}^L (L - z_L) \tag{7.3.23}$$

This leads to

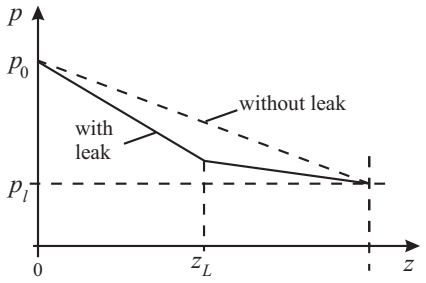


Fig. 7.8. Pressure profile in a horizontal pipeline before and after occurrence of a leak

$$z_L = \frac{p_l - p_0 + p_{zII}^L l}{p_{zII}^L - p_{zI}^L} = l \frac{p_{zII}^L - p_z^L}{p_{zII}^L - p_{zI}^L} \tag{7.3.24}$$

Introducing small changes of the gradients after a leak

$$\left.\begin{aligned} \Delta p_{zI} &= p_{zI}^L - p_z \\ \Delta p_{zII} &= p_{zII}^L - p_z \end{aligned}\right\} \tag{7.3.25}$$

yields for the leak location

$$z_L = l \frac{\Delta p_{zII}^L}{\Delta p_{zII}^L - \Delta p_{zI}^L} \tag{7.3.26}$$

In order to filter disturbances the upstream and downstream pressure gradients are estimated by recursive averaging:

$$p_{zj}(k) = \kappa_p p_{zj}(k-1) + (1 - \kappa_p) \frac{\dot{m}_j^2(k)}{p_0(k) - p_l(k)} \quad (j = I, II) \tag{7.3.27}$$

Then reference mass flows are determined

$$\dot{m}_j^{*2}(k) = p_{zj}(k)[p_0(k) - p_l(k)] \quad (j = I, II) \tag{7.3.28}$$

As soon as a leakage alarm is triggered the pressure gradients $p_{zj}(k)$ are no longer calculated but fixed and the mass flow differences

$$\Delta m_j^2(k) = \dot{m}_j^2(k) - \dot{m}_j^{*2}(k) \quad (j = I, II) \tag{7.3.29}$$

are determined and averaged:

$$\overline{\Delta m_j^2(k)} = \frac{1}{N}\sum_{k=1}^{N}\Delta m_j^2(k) \quad (j = I, II) \tag{7.3.30}$$

In stationary operation (7.2.15) reduces to

$$\begin{aligned} \frac{\partial p}{\partial z} &= -k_F \dot{m}^2 \\ k_F &= \frac{\lambda}{2 d_F A_F^2 \rho} = \frac{\lambda c_F^2(p)}{2 d_F A_F^2 p} \end{aligned} \tag{7.3.31}$$

Therefore (7.3.21) and (7.3.22) can be expressed by

$$\left.\begin{aligned} \Delta p_{zI}^L &= p_{zI}^L - p_z = k_F(\dot{m}_0 + \Delta\dot{m}_0)^2 \\ \Delta p_{zII}^L &= p_{zII}^L - p_z = k_F(\dot{m}_l + \Delta\dot{m}_l)^2 \end{aligned}\right\} \tag{7.3.32}$$

Neglecting small terms $\Delta\dot{m}^2$, setting $\dot{m}_0 = \dot{m}_l$ and inserting in (7.3.26) yields for the *leak location*

$$\hat{z}_L = l\frac{1}{1 - \frac{\overline{\Delta\dot{m}_0}}{\overline{\Delta\dot{m}_l}}} \tag{7.3.33}$$

which allows one to calculate the leak location only from mass flow rate changes determined by (7.3.30). An alternative with autocorrelation functions

$$R_{m_0 m_0}(\tau) = \frac{1}{N}\sum_{k=1}^{N}\Delta\dot{m}_0(k-\tau)\Delta\dot{m}_0(k)$$

$$R_{m_l m_l}(\tau) = \frac{1}{N}\sum_{k=1}^{N}\Delta\dot{m}_l(k-\tau)\Delta\dot{m}_l(k)$$

is

$$\hat{z}_L = l\frac{1}{1 - \frac{R_{\sum m_0}}{R_{\sum m_l}}} \tag{7.3.34}$$

with autocorrelation function averages like (7.3.17).

After the appearance of a leak mass flow rate it holds that

$$\dot{m}_L = \left(\dot{m}_0 + \overline{\Delta\dot{m}_0}\right) - \left(\dot{m}_l + \overline{\Delta\dot{m}_l}\right) = \overline{\Delta\dot{m}_0} - \overline{\Delta\dot{m}_l} \tag{7.3.35}$$

Hence, it is $\Delta\dot{m}_0 > 0$ and $\Delta\dot{m}_l < 0$ and the ratio z_L/l in (7.3.34) becomes

$$z_L/l < 1$$

The *leak mass flow rate* can be calculated with (7.3.35).

7.3.3 Leak detection for gas pipelines

a) Gas pipeline models

Different to liquid pipelines the state equation cannot be linearized for gas pipelines, because the fluid is compressible and the parameters become dependent on the pipeline section j. Beginning with the "long pipeline model" (7.2.14) and (7.2.15), the basic mass balance and momentum balance equation becomes

$$\frac{A_F}{c_F^2}\frac{\partial p}{\partial t} + \frac{\partial \dot{m}}{\partial z} = 0 \tag{7.3.36}$$

$$\begin{aligned}\frac{1}{A_F}\frac{\partial \dot{m}}{\partial t} + \frac{\partial p}{\partial z} &= -\frac{\lambda c_F^2}{2 d_F A_F^2}\frac{\dot{m}|\dot{m}|}{p} - \frac{g \sin\alpha}{c_F^2} p \\ &= -F - Y\end{aligned} \tag{7.3.37}$$

These equations are identical with (7.2.14), (7.2.15) but contain the static pressure dependent on the angle α of the height profile gradient.

If the frequently made assumption that the speed of sound c_F is constant is not acceptable, one can approximate the dependence by a linear relation, [7.2]:

$$c_F(p) = \overline{c}_F + \frac{\partial c_F}{\partial p}(p - \overline{p}) \tag{7.3.38}$$

with $\partial c_F / \partial p = \text{const}$. Then for the mass balance equation it follows that

$$\frac{\partial}{\partial t}\left(\frac{p}{c_F^2(p)}\right) = \frac{1}{\overline{c}_F^2}\left(1 - \frac{2\overline{p}}{\overline{c}_F}\frac{\partial c_F}{\partial p}\right) = \frac{1}{c_F^2}\beta$$

and therefore with the correction factor $\beta = \text{const}$.

$$\beta\frac{A_F}{c_F^2}\frac{\partial p}{\partial t} + \frac{\partial \dot{m}}{\partial z} = 0 \tag{7.3.39}$$

For the momentum balance equation (7.3.37) just $c_F(p)$ from (7.3.38) has to be inserted. Then the following equation system results:

$$\begin{bmatrix} \beta\frac{A_F}{\overline{c}_p^2} & 0 \\ 0 & \frac{1}{A_F} \end{bmatrix}\begin{bmatrix} \frac{\partial p}{\partial t} \\ \frac{\partial \dot{m}}{\partial t} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{\partial p}{\partial z} \\ \frac{\partial \dot{m}}{\partial z} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{\lambda|\dot{m}|\dot{m}c_F^2(p)}{2 d_F A_F^2 p} - \frac{p g \sin\alpha}{c_F^2(p)} \end{bmatrix} \tag{7.3.40}$$

This equation system is now solved numerically for time intervals $t = k\Delta t$ and N pipeline sections:

$$\Delta z = \frac{l}{N} \tag{7.3.41}$$

as shown in Figure 7.9. As an approximation for the derivatives a centered difference scheme is introduced, [7.2], [7.4]:

$$\left.\frac{\partial x}{\partial t}\right|_{z,k} = \frac{3x_z^{k+1} - 4x_z^k + x_z^{k-1}}{2\Delta t}$$
$$\left.\frac{\partial x}{\partial z}\right|_{z,k} = \frac{x_{z+1}^{k+1} - x_{z-1}^{k+1} + x_{z+1}^k - x_{z-1}^k}{4\Delta z} \tag{7.3.42}$$

Then the following equation system results:

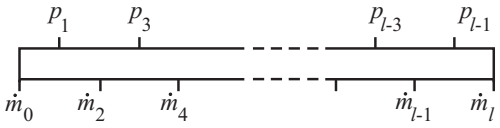


Fig. 7.9. Discretization of the pipeline for numerical simulation

$$\mathbf{A}\mathbf{x}^{k+1} = \mathbf{f}\left(\mathbf{x}^k, \mathbf{x}^{k-1}\right) + \mathbf{s}\left(p_0^{k+1}, p_N^{k+1}\right) \tag{7.3.43}$$

with the state vector

$$\mathbf{x}^k = \left[\dot{m}_0^k, \dot{m}_2^k, \ldots, \dot{m}_l^k, p_1^k, p_3^k, \ldots, p_{l-1}^k\right]^T \tag{7.3.44}$$

As the system matrix is constant, the linear equation system (7.3.43) is solved by the following state equation for the simulation of the pipeline variables:

$$\mathbf{x}^{k+1} = \mathbf{A}^{-1}\left[\mathbf{f}(\mathbf{x}^k, \mathbf{x}^{k-1}) + \mathbf{s}(p_0^{k+1}, p_N^{k+1})\right]$$

and output equation for the flows

$$\mathbf{y}^{k+1} = \begin{bmatrix} \dot{m}_0^{k+1} \\ \dot{m}_l^{k+1} \end{bmatrix} = [1, 0, \ldots, 1, 0, \ldots, 0]\ \mathbf{x}^{k+1} \tag{7.3.45}$$

This solution needs only a small computational effort by multiplying the inverse matrix with a vector including a nonlinear function of the two last states, the friction coefficient, the height correction and the two input pressure signals p_0 and p_N. The output signals $\dot{m}_0$ and $\dot{m}_N$ are elements of the state vector. Tests with more complex solution methods (nonlinear, but large computational effort) gave rather similar results for the considered application. For more details see [7.2].

b) Leak detection with state reconstruction and correlation functions

It is assumed that a small leak flow rate $\dot{m}_L$ occurs at location z_L. This effect is taken into account by introducing this loss in the mass balance for the influenced section. Therefore an enlarged pipeline model results with the leak influence vector I, dependent on the leak location:

$$\mathbf{x}^{k+1} = \mathbf{A}^{-1}\left[\mathbf{f}\left(\mathbf{x}^{k}, \mathbf{x}^{k-1}\right) + \mathbf{s}\left(p_0^{k+1}, p_N^{k+1}\right)\right] + \mathbf{l}\dot{m}_L \tag{7.3.46}$$

In order to detect leaks also for wide ranges of operating conditions *nonlinear pipeline models* have to be used. This leads to nonlinear state observers. An additional requirement is that the information on the leak should not vanish with time. For the pipeline model (7.3.45) most of the coefficients are known with good accuracy, except the friction coefficient λ which may also change with time. Therefore, this coefficient will be estimated (online) by the least-squares method. This leads to an *adaptive (nonlinear) state observer*. An advantage of this approach is furthermore that the estimated friction coefficient does not change the steady-state solution of the mass balance in (7.3.37) so that leak effects will not be compensated by the observer.

Figure 7.10 shows the resulting leak supervision structure including the pipeline observer and the leak detection monitor, where both differences x and y act as residuals.

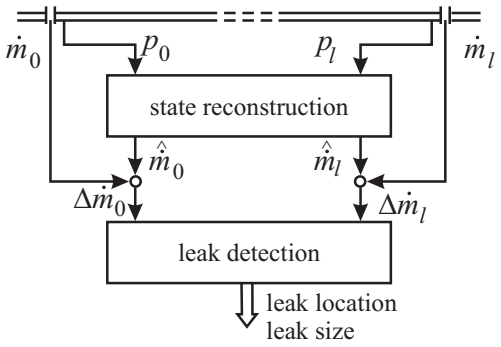


Fig. 7.10. Leak detection for gas pipelines with state reconstruction

The corresponding equations are:

- pipeline:

$$\begin{aligned}\mathbf{x}^{k+1} &= \mathbf{A}^{-1}\left[\mathbf{f}\left(\mathbf{x}^{k}, \mathbf{x}^{k-1}, \lambda, \mathbf{h}\right) + \mathbf{s}\left(p_0^{k+1}, p_l^{k+1}\right)\right] + \mathbf{l}\dot{m}_L \\ \mathbf{y}^{k+1} &= [1, 0, \ldots, 1, 0, \ldots, 0]\,\mathbf{x}^{k+1}\end{aligned} \tag{7.3.47}$$

- observer:

$$\begin{aligned}\hat{\mathbf{x}}^{k+1} &= \mathbf{A}^{-1}\left[\mathbf{f}\left(\hat{\mathbf{x}}^{k}, \hat{\mathbf{x}}^{k-1}\lambda, \mathbf{h}\right) + \mathbf{s}\left(p_0^{k+1}, p_l^{k+1}\right)\right] \\ \hat{\mathbf{y}}^{k+1} &= [1, 0, \ldots, 1, 0, \ldots, 0]\,\hat{\mathbf{x}}^{k+1}\end{aligned} \tag{7.3.48}$$

- residuals:

$$\mathbf{e}^{k} = \mathbf{y}^{k} - \hat{\mathbf{y}}^{k} = \begin{bmatrix} \dot{m}_0^k - \hat{\dot{m}}_0^k \\ \dot{m}_l^k - \hat{\dot{m}}_l^k \end{bmatrix} = \begin{bmatrix} \Delta\dot{m}_0^k \\ \Delta\dot{m}_l^k \end{bmatrix} \tag{7.3.49}$$

To show the effects of a suddenly appearing leak, a gas pipeline was simulated with $\dot{m}_L = 0.35\,\mathrm{kg\,s^{-1}}$ and $z_L/L_R = 0.5$. More data are given in Section 7.4.2.

The mass flow rates $\Delta\dot{m}_0$ at the beginning and $\Delta\dot{m}_l$ at the end of the pipeline change in predetermined directions depending on the leak flow rate and the leak location, see Figure 7.11.

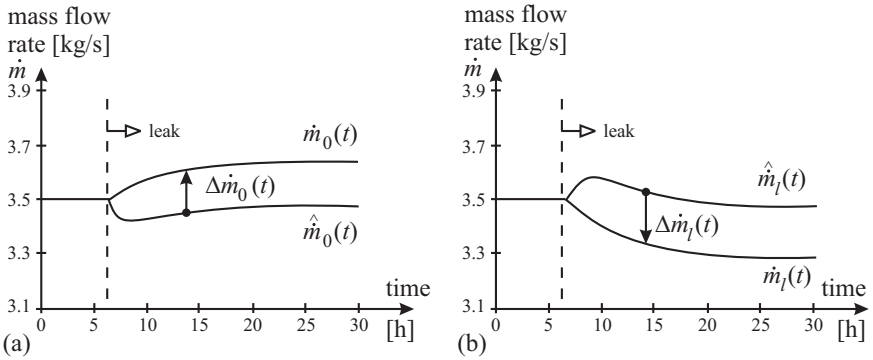


Fig. 7.11. Changes of the gas mass flow rate after a leak: a) at the beginning; b) at the end of a pipeline. Simulation for $z_L/l = 0.5$ with $\dot{m}_L = 0.35\,\mathrm{kg\,s^{-1}}$ (8%)

A sensitive decision algorithm for "leak" or "no leak" is the cross-correlation function, compare (7.3.16),

$$R_{MM}(\tau) = E\{\Delta\dot{m}_0(k-\tau)\Delta\dot{m}_l(k)\} \tag{7.3.50}$$

which results (theoretically) in

$$R_{MM}(\tau) = \begin{cases} 0 & \text{no leak} \\ -f(\dot{m}_L, z_L) & \text{with leak} \end{cases} \tag{7.3.51}$$

which means it changes in a predetermined direction. The computation is realized by a recursive filter of first order as (7.3.19).

To reduce noise effects the alarm criterion is taken as the sum over several time shifts τ:

$$R_\Sigma = \frac{1}{2M+1} \sum_{\tau=-M}^{M} R_{MM}(\tau) \tag{7.3.52}$$

This cross-correlation sum reacts sensitively even to small leaks. An alarm is given when the sum crosses a predefined alarm threshold.

After a leak is detected, the parameter estimation of λ is frozen and the estimation of the leak location starts. Introducing the auto-correlation sums as for (7.3.34), the leak location is, according to (7.3.34), estimated by

$$\hat{z}_L(k) = \frac{l}{1 - \frac{R_{\Sigma m_0}}{R_{\Sigma m_l}}} \tag{7.3.53}$$

where l_R is the length of the pipeline, [7.12]. The leak flow rate is estimated by the dynamic balance equation

$$\dot{m}_L(k) = \Delta\dot{m}_0^k - \Delta\dot{m}_l^k \tag{7.3.54}$$

compare (7.3.35).

Tests will be described which are based on measured signals of a gas pipeline. In order to simulate leaks, the observer is enlarged by the leak influence vector I with a negative leak flow rate. This realization (also useful for self-test) respects the dynamics of the leak influence and is a good approximation for a real suddenly appearing leak in the pipeline according to the resulting residual, [7.1].

The ethylene pipeline is 150 km long with a diameter of 0.26 m and a varying speed of sound (depending on the pressure). The four measured signals ($\dot{m}_0$, p_0, $\dot{m}_l$, p_l) were sampled every 3 minutes. The observer uses a time interval of 30 s, so that the measured signals have to be interpolated. With the section length interval of about 9.4 km the system order is 17. Furthermore, an approximation of the geographic height profile is included, see Figure 7.12.

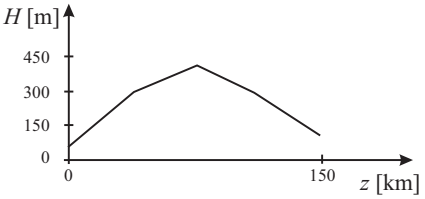


Fig. 7.12. Approximation of the geographic height profile of an ethylene pipeline

The decision algorithm is modified, substituting $\Delta\dot{m}_0(k)$ and $\Delta\dot{m}_N(k)$ in (7.3.19) and (7.3.50) by

$$\begin{aligned} \Delta\dot{m}_0'(k) &= \Delta\dot{m}_0(k) - E\{\Delta\dot{m}_l(k)\} \\ \Delta\dot{m}_l'(k) &= \Delta\dot{m}_l(k) - E\{\Delta\dot{m}_0(k)\} \end{aligned} \tag{7.3.55}$$

This yields a better sensitivity and is independent of the leak location.

Figure 7.13 shows the measured input and output pressure of the described pipeline during a test period of 65 h.

As illustrated in Figures 7.14a) and 7.14b), the pipeline observer describes the dynamic behavior of the pipeline quite well, so that a sensitive leak detection is possible. Results from several leak simulations with different leak ratios (leak ratio given relative to a mean flow rate) are given in Figure 7.15. Figure 7.16 shows an example of the leak location estimation for a leak ratio of 5%.

Furthermore, the time for computation (less than 2 s for a PDP11/34) is rather small in comparison to the sampling time of 3 min.

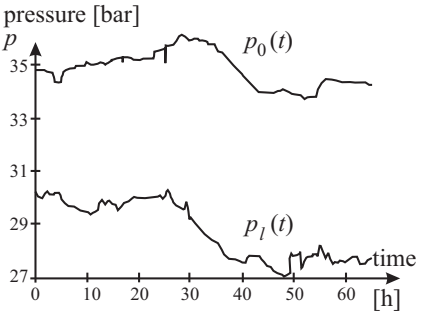


Fig. 7.13. Measured input and output pressure of the ethylene pipeline

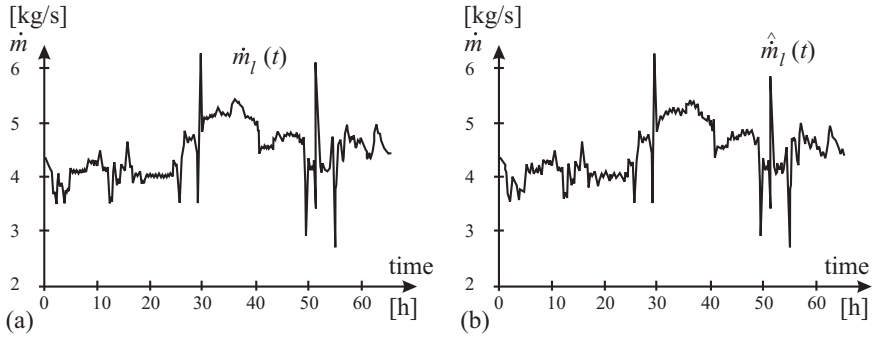


Fig. 7.14. Outlet flow rates for the gas pipeline: a) measured; b) reconstructed

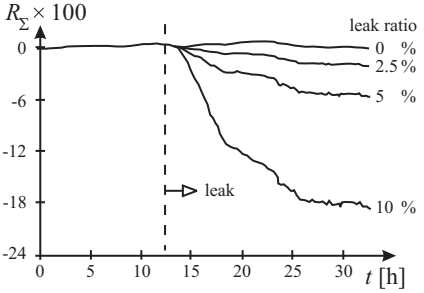


Fig. 7.15. Development of the averaged cross-correlation detector in dependence on the size of a leak at $z_L = 80\,\text{km}$

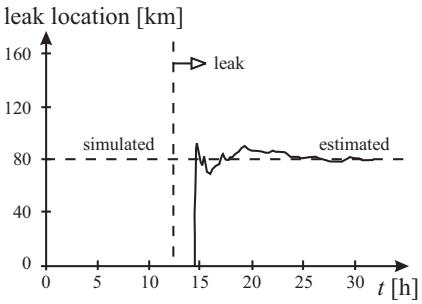


Fig. 7.16. Development of the leak location detector for a leak of 5%

7.4 Experimental results

7.4.1 Gasoline pipeline

The mass-balanced leak detection method with correlation analysis described in Section 7.3.2 was experimentally tested on a gasoline pipeline of $l = 48$ km and $d_F = 273$ mm. Figure 7.17 shows the course of the pipeline considering the height above sea level as well as the location of pumps. The two main pumps are driven by 400 kW asynchronous machines, which may be operated individually or together. At full power about 330 $m^3\ h^{-1}$ are delivered at an initial pressure of $p_{in} = 69$ bar. This pressure is measured after the pumps, but before the entrance valve. The line has a

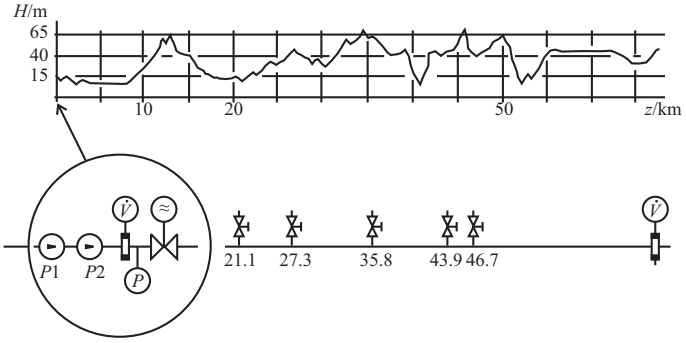


Fig. 7.17. Scheme of the gasoline pipeline with topographic profile and measurements

wall thickness of 8 mm. Intermediate depots are located at 21.1, 27.3, 35.8, 43.9 and 46.7 km.

The volume flows are measured by means of measuring orifices and Barton cells, and pressures with Barton cells (accuracy about 0.1%). The volume flow $\dot{V}_l$ at the end of the line is transmitted by a telemetric device, i.e. deviations from the operating point – about 1/10 of the total measuring range (0–400 $m^3\ h^{-1}$) – are encoded into

an 8-bit word. This corresponds to a resolution of 0.16 $m^3\ h^{-1}$ or 0.05% with respect to 330 $m^3\ h^{-1}$.

Since the pressure at the end of the line was almost constantly equal to atmospheric pressure, recording and processing of the measurement variable p_l was neglected. Thus only both volumetric flows $\dot{V}_0$ and $\dot{V}_l$, as well as the pressure at the beginning of the pipeline p_{in}, were used for leakage monitoring. For further details on the pipeline and the equipment see [7.13].

An Intel MDS800 microcomputer development system was used to carry out the online experiments. An 8-bit microcomputer with the 8080 A central processor was used. The system is extended to 48 k of RAM.

A program package – 16 kbytes of program memory – for leakage monitoring was implemented on the microcomputer system in the ASM 80 assembler language.

A series of experiments for leakage detection were carried out on the pipeline, where leaks could be generated artificially at the branches to the intermediate depots.

The following values were assumed for the constants:

$$\begin{aligned} \kappa_m &= 0.0075 \\ \lambda &= 0.99 \\ R_{\Sigma th} &= -0.5[\mathrm{m}^3/\mathrm{h}]^2 \\ P &= 20 \end{aligned}$$

Figure 7.18 shows one of the experiments carried out. The values of the signals p_{in}, $\dot{V}_0$ and $\dot{V}_l$ collected by the microcomputer were recorded. The sum of cross-correlation functions R_Σ for the indication of a leakage is shown. The leak was generated at $t = t_L$. The leakage location and leakage flow calculation after detection of the leak is also shown.

For the experiment shown – leakage location at 35.8 km and mean leakage rate of 0.19% which is about 0.2 l/s – the alarm trigger level was exceeded 98 s after occurrence of the leak. The leak location was estimated with an error of about ±0.7% or ±500 m at time 90 s after the alarm.

Since the characteristic variable R_Σ in all cases with leaks significantly exceeded its standard deviation during regular operation without a leak, it may be assumed that even smaller leaks than those in the experiments carried out can be detected and localized.

A comparison with low-pass-filtered mass balance with the same measured data has shown that the thresholds must be considerably wider (about 3 to 4 times). More details are given in [7.11].

7.4.2 Gas pipeline

As it is very difficult to generate real leaks in gas pipelines (large volume, no available storage, loss of valuable products, environmental problem) only simulated leaks, but with measured pipeline measurements are shown. The leak is introduced via the leak influence vector **l** of (7.3.46) abruptly, to simulate a suddenly appearing leak

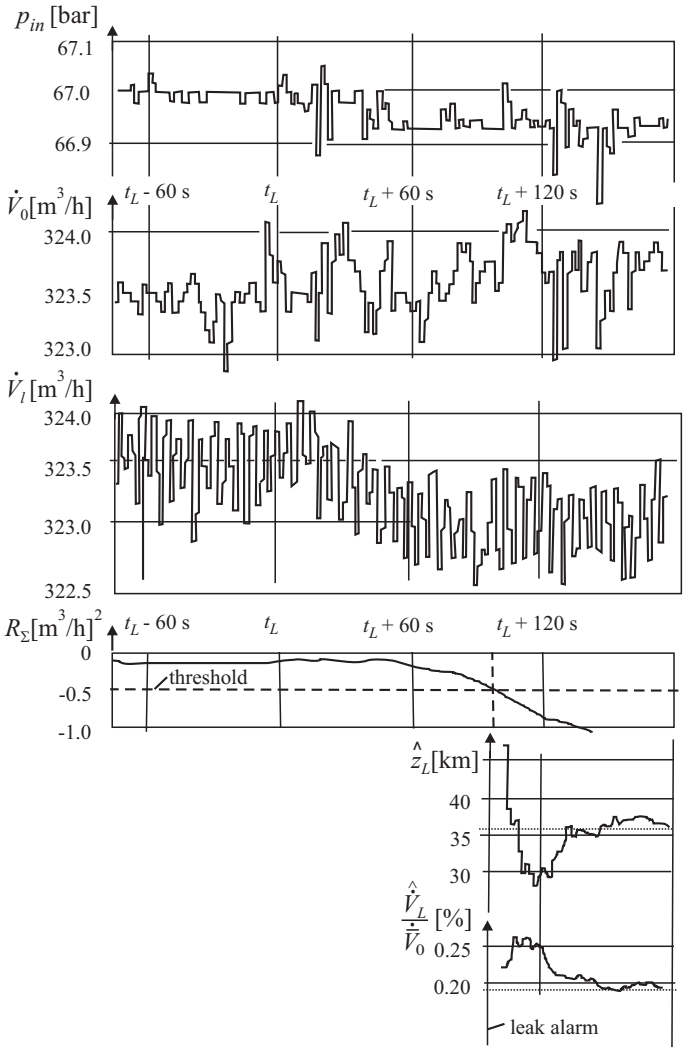


Fig. 7.18. Measured variables $p_{in}(t)$ and $\dot{V}(t)$ of the gasoline pipeline before and after the leak of 0.19% at $l = 35.8$ km. R_Σ: cross-correlation, $\hat{z}_L$: leak location estimate, $\hat{\dot{V}}_L$: leak flow estimate

(e.g. crack of welding seam). The pipeline transports ethylene, is l=150 km long with inner diameter $l_F = 0.26$ m. The sampling time of the four measured signals $\dot{m}_0, p_0, \dot{m}_l, p_l$ is $T_0 = 3$ min. For the simulated state reconstruction the length of each section is 9.4 km, such that 17 sections result. The sampling time for the model is $T_{0,m} = 30$ s. Therefore measured signals are interpolated. The height profile is shown in Figure 7.12 and measured pressure signals over 65 h in Figure 7.13. The speed of sound was assumed as dependent on the pressure and is about 260 m/s. The computation time for the leak detection algorithms is around 1 s.

The agreement of measured reconstructed flow rates is relatively good, Figure 7.14, compared to the resolution of the telemetry system and noise effects. The time history of the cross-correlation function (7.3.17), Figure 7.15 indicates that leaks of $\geq 2\%$ can be detected and that detection time is the smaller the larger the leak. The detection time for a leak of 2% is about 10 h and for 5% about 3.5 h. The leak location estimation reacts relatively fast. A first estimation is for a 5% leak available after about 2 h and after 15 h the location is met relatively precisely. Also leak size estimation needs about 15 h with an accuracy of about 8%, Figure 7.16. These results are significantly better than detection by mass flow balance, where only leaks larger than about 10% can be detected.

7.4.3 Conclusions

Simulations, measurements and leak experiments have shown that the early detection and localization of small leaks in liquid and gas pipelines can be considerably improved by model-based methods. The leak-detection methods are based on mathematical dynamic models, nonlinear adaptive state reconstruction and a correlation detection technique. The measured signals are one flow rate and one pressure at each end of a pipeline. As the required computational effort is relatively small, microcomputers or personal computers can be used.

Commercialized versions of the described methods and further adaptation to special liquid pipelines with applications to several pipelines in Germany, Austria and Russia are reported in [7.10].

8

Fault diagnosis of industrial robots

This chapter describes firstly how analytical symptoms can be obtained by the estimation of physical defined parameters for the six axes of a multi-axis industrial robot. Then it is shown how the analytic symptom knowledge is added by heuristic symptoms observed by the maintenance personnel as described in Section 2.3 and Figure 2.7 and how a fault diagnosis can be performed with both the analytic and heuristic symptom information by using fuzzy-logic inferencing. As industrial robots (IR) are usually servo systems with point-to-point movements or trajectory following they have sufficient dynamic excitation and therefore parameter estimation can be preferably applied for fault detection.

8.1 Structure of a six-axis robot

The application of a knowledge-based fault-diagnosis strategy is described for an industrial robot of type Jungheinrich R106, see Figure 8.1. The device consists of six revolving joints actuated by DC servomotors of high dynamic performance. The following considerations concentrate on the investigation of the mechanical subsystem of the different axes because a strong demand for preventive maintenance and incipient fault diagnosis exists for this part of the robot, [8.1], [8.2], [8.5].

The mechanical drive chains of the axes consist of different standard machinery elements (gears, bearings, toothed belts, shafts, etc.), transferring torque from the motor to the moved (actuated) arm as shown in Figure 8.2.

The control of each axis is performed by a cascaded control with an inner speed control of the DC motor and an outer position control of the axis joint. Figure 8.3 depicts the signal flow. The measured variables for the fault detection are:

φ joint position
ω motor speed
I_A armature current of DC motor

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_8, © Springer-Verlag Berlin Heidelberg 2011

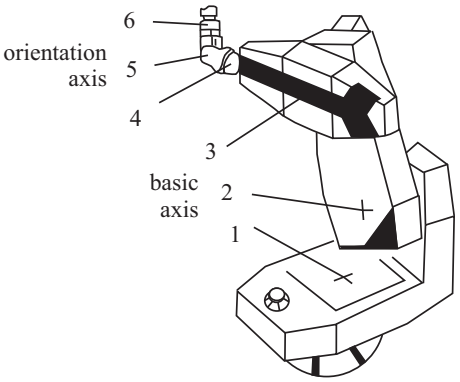


Fig. 8.1. Industrial robot with 6 rotational axes

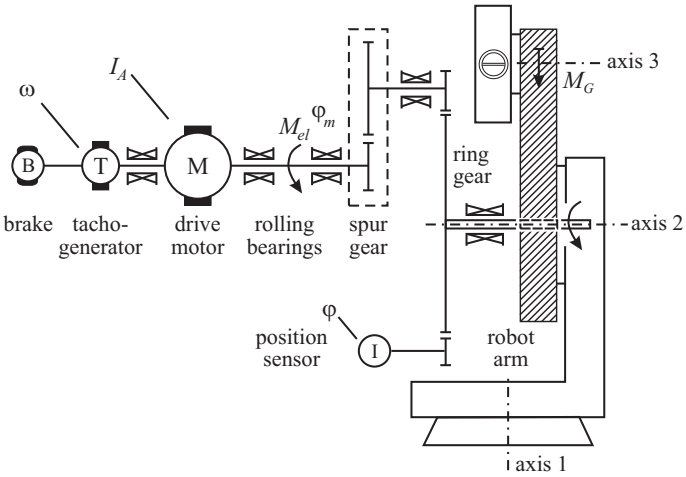


Fig. 8.2. Scheme of the mechanical drive chain (axis) of an industrial robot. Measurable quantities are φ, ω, I_A

8.2 Model of a robot axis and parameter estimation

Assuming the arms as rigid bodies each joint can be modeled by stating a torque balance related to the joint axis

$$M_{el}(t)/\nu_i = J_L(\varphi_0, m_L)\ddot{\varphi}(t) + M'_{F0}\,\mathrm{sign}\,\dot{\varphi}(t) + M'_{F1}\dot{\varphi}(t) + M'_G(m_L, \varphi_0) \quad (8.2.1)$$

where

$M_{el} = \Psi_A I_A$	electrical torque at motor output axle
Ψ_A	armature flux linkage
I_A	armature current
ν	total gear ratio φ/φ_m

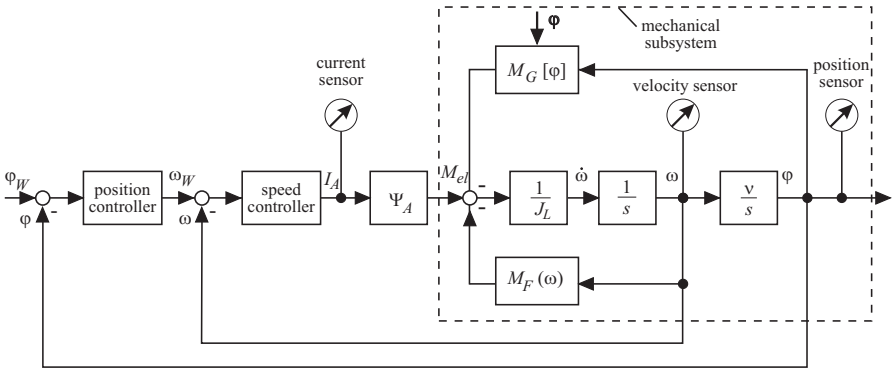


Fig. 8.3. Block diagram of the model of an industrial robot drive unit with conventional cascaded closed control

J_L	moment of inertia of the arm (position and load dependent)
M'_{F0}	Coulomb-friction torque on joint side
M'_{F1}	viscous-friction torque on joint side
M'_G	gravitational torque on joint side
m_L	mass of load at end effector
φ	arm position
φ_0	arm base position
$\omega = \dot{\varphi}/\nu$	motor angular speed

The gravitation torque is modeled by

$$M_G(m_L, \varphi_0) = M'_{G0} \cos \varphi \tag{8.2.2}$$

and may be dependent on a kinematic gravitational torque compensation device, e.g. a pneumatic cylinder. The couplings between the axes can be neglected if the movements are not very fast.

Ψ_A is known from the motor´s data sheet. Discretizing the continuous-time model (8.2.1) with $k = t/T_0$, T_0 sampling time, and relating the parameters to the motor side by multiplying by ν leads to

$$M_{el}(k) = J(\varphi_0, m_L)\dot{\omega}(k) + M_{F0} \operatorname{sign} \omega(k) + M_{F1}\omega(k) + M_{G0} \cos \varphi(k) \tag{8.2.3}$$

($1/\nu$ is for axes 1–6: 197, 197, 131, 185, 222, 194)

Then this equation results in vector notation:

$$\begin{aligned} M_{el}(k) &= \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\Theta}}(k) + e(k) \\ \boldsymbol{\psi}^T(k) &= [\dot{\omega}(k),\ \operatorname{sign} \omega(k), \omega(k) \cos \varphi(k)] \\ \hat{\boldsymbol{\Theta}} &= \left[\hat{J}, \hat{M}_{F0}, \hat{M}_{F1}, \hat{M}_{G0}\right] \end{aligned} \tag{8.2.4}$$

and is used for recursive parameter estimation in continuous time, see [8.3], with

$$\dot{\omega}(k) = \left.\frac{d\omega(t)}{dt}\right|_k = \frac{\omega(k) - \omega(k-1)}{T_0}$$

where T_0 is a small sampling time and $e(k)$ an equation error. Note that here the estimated process parameters are identical to the physically defined process coefficients.

8.3 Analytic and heuristic diagnosis knowledge

As discussed in Section 2.3, see Figure 2.7, the overall diagnosis of many technical processes is based on analytic and heuristic information on the process state. The example of the industrial robot is now used to show how the results of both information sources can be implemented and how conclusions for a final diagnosis can be drawn, [8.1]. The diagnosis procedure follows Chapters 15, 16 and 17 of [8.4].

8.3.1 Symptom representation

The fault diagnosis is based on symptoms available from different sources. The following categories are defined.

a) Analytical symptoms

An analytically determined set of symptoms S_a is stored in a specific predefined area in process computer memory. This "symptom buffer" is arranged as follows:

< record SYMPTOMS S_{ia} of >

(1) Number of robot axes the symptom is related to
(2) Symptom name expressed as a symbolic string
(3) Numerical mean value of symptom (coefficient)
(4) Numerical nominal value of symptom (coefficient)
(5) Physical unit of symptom (coefficient)
(6) Calculated confidence number $\mu(S_{ia})$
(7) Time of symptom entry into buffer
(8) Explanatory text concerning the specific symptom.

At the time of diagnosis (specific time intervals or upon operator´s request) these analytically calculated symptoms (represented by deviations of process coefficients from their nominal values) are stored in this buffer. It therefore acts as an interface between the analytical and the heuristic part of the knowledge-based robot fault-diagnosis system.

b) Heuristic symptoms

The second class of symptoms is represented by a set of heuristic symptoms S_b, neither directly measurable nor analytically computable, e.g. empirical knowledge of the maintenance personnel operating on the manufacturing floor. These symptoms have to be provided in an interactive dialogue with the diagnosis system.

c) Process history and fault statistics

A third category of facts, frequently available in the industrial environment, depends on the general status of the diagnosed industrial robot. For consistency reasons concerning the arrangement of the knowledge base, these facts are treated in the same way as the symptoms. Corresponding data from mass storage files can be treated as symptoms belonging to class a), while specific heuristic knowledge according to maintenance details is considered part of class b).

8.3.2 Diagnosis knowledge representation

As a suitable tool for the systematic treatment of heuristic knowledge (especially in the diagnostic domain) inference methods can be applied in order to set up logical interactions between observed symptoms (effects) and unknown faults (causes) and to structure the knowledge in a problem-adapted manner. A single rule is described by the expression

$$\text{IF} < \text{condition} > \text{THEN} < \text{conclusion} > \tag{8.3.1}$$

where the condition part contains facts (symptoms) as inputs being associated by boolean AND and OR connectives. The conclusion part represents a so-called "event" as a logical cause of these facts. Chaining of the rules now establishes the causal dependencies of symptoms and faults (considered as "basic events" in a hierarchical manner). Thus intermediate events $E_k, k = 1, ..., j$ are introduced. This natural procedure results in the establishment of a fault-symptom tree ("directed graphic") structuring the rules hierarchically, relating symptoms to events and faults by a systematic approach, see Section 2.6.2 and [8.4], Chapter 17.

A systematic treatment of analytical and heuristic symptoms requires a unified symptom representation. As described in [8.4], Chapter 15, this can be based on confidence numbers $0 \leq c(S_i) \leq 1$ for the symptoms S_i or on membership functions $0 \leq \mu(S_i) \leq 1$ in the frame of fuzzy logic.

The diagnostic reasoning with inference methods is usually based on forward and backward chaining. By using *forward chaining* for a rule like (8.3.1) the facts are matched with the premise and the conclusion is drawn based on the logical consequence (modus ponens). The condition part (premise) contains facts in the form of symptoms S_i as inputs, and the conclusion part includes events E_k and faults F_j as a logical cause of the facts. If several symptoms indicate an event or fault, the facts are associated by AND and OR connectives, leading to rules like

$$\text{IF } < S_1 \text{ AND } S_2 > \text{ THEN } < E_1 > \tag{8.3.2}$$

$$\text{IF } < E_1 \text{ OR } E_2 > \text{ THEN } < F_1 > \tag{8.3.3}$$

As the symptoms usually have to be considered as uncertain facts the approximate reasoning with fuzzy logic is an appropriate way of treating the causalities in a systemic and unified way.

Therefore the facts are assigned to fuzzy sets with membership functions $\mu(S_{ia}$ and $\mu(S_{ih})$. The approximate reasoning with fuzzy logic and a simplified version with singletons as outputs is described in [8.4], Chapter 17.

The conditional part of the IF-THEN rules can be evaluated by the max-min composition to obtain the most possible fault:

$$\text{Fuzzy AND} : \mu(\eta) = \min[\mu(\xi_1), \ldots, \mu(\xi_\nu)] \tag{8.3.4}$$

$$\text{Fuzzy OR} : \mu(\eta) = \max[\mu(\xi_1), \ldots, \mu(\xi_\nu)] \tag{8.3.5}$$

An alternative is the prod-sum operation. The NOT operation follows:

$$\text{NOT} : \mu(\eta) = 1 - \mu(\xi) \tag{8.3.6}$$

The strategy of *backward chaining* assumes the conclusion as known and searches for the relevant premises (modus tollens). This is especially of interest if the symptoms are not complete. The concluded events and faults are then displayed to the operator after forward chaining with all known symptoms. However, this requires an interactive dialogue with the operator.

8.3.3 Faults, heuristic symptoms and events of the robot

The analytical and heuristic symptoms are now considered for the investigated industrial robot. According to maintenance manuals and discussions with IR-manufacturers the following list of *possible faults* in the mechanical components of robot drives results:

F_1 Loosened bracing screw of spur gearing
F_2 Significant wear in axis drive mechanics
F_3 Bracing in the drive chain too high
F_4 Overheating of motor-gear unit
F_5 Overload of axis drive by payload or tool
F_6 Defect in electromagnetic brake (incomplete release)

The parameter estimation yields the following *analytical symptoms*:

S_{1a} Decrease of J_L
S_{2a} Increase of J_L
S_{3a} Decrease of M_{F0}
S_{4a} Increase of M_{F0}
S_{5a} Decrease of M_{F1}
S_{6a} Increase of M_{F1}

From many experimental investigations with the robot, data from the manufacturer and knowledge of the specific device, the following non-measurable *heuristic symptoms* result, including specific symbolic names:

S_{1h} Characteristic acoustic noise I present $\rightarrow$ char_noise_I
S_{2h} Characteristic acoustic noise II present $\rightarrow$ char_noise_II

S_{3h} Obvious inaccuracies in positioning → inacc_pos
S_{4h} Manual test: movement of axis sluggish → test_move
S_{5h} Assumption: backlash in drive chain → backl_assum

It is assumed that the following facts are available from *process history* and *fault statistics*:

S_1 Last maintenance: short/long time ago → last_maint
S_2 Number of operating hours: low/high → operat_hours
S_3 Mechanical collision (possibly) occurred → mech_coll

Figure 8.4 outlines an extract of the implemented knowledge base in the form of multilevel fault-symptom trees according to the specific faults F_1, F_2 and F_3 indicating causal relationships. Intermediate steps of diagnosis are included in the directed graphs by *events* representing effects of different faults.

E_1 Bracing in drive chain actually too low → brace_low
E_2 Significant decrease of friction in drive chain → fric_decr
E_3 Sluggish mechanical movement in axis → slugg_mov
E_4 Significant increase in friction of drive chain → fric_incr
E_5 Increase of backlash in drive chain → b_lash_incr
E_6 Limit cycles of specific axis → limit_cycles

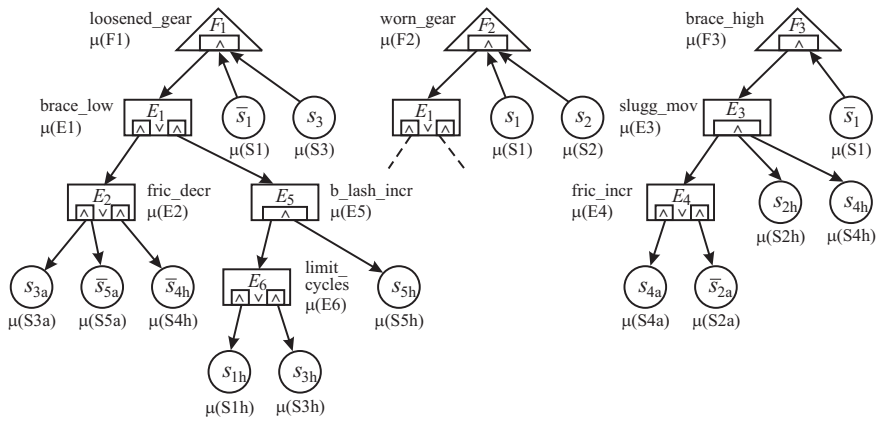


Fig. 8.4. Extract of the analytical and heuristic knowledge base (fault-symptom trees) of the industrial robot diagnosis system for one axis ($\bar{S}$ means < NOT S >)

8.4 Experimental results

The fault-diagnosis procedure will now be shown exemplarily for basic axis 1 in the case of faults F_1, F_2 and F_3. Defects F_1 and F_3 can originate, e.g. from erro-

neous assembly/maintenance or from mechanical collision (short or medium operating hours). F_2 is mainly the result of wear in the drive mechanics (long operating hours).

8.4.1 Fault diagnosis with analytical knowledge

Figure 8.5 illustrates the typical behavior of the measured signals in the case of point-to-point movement of basic axis 1. The end effector did not carry any extra load. The sampling interval $T_0 = 5\,\text{ms}$ is identical to that of the position controller for embedding the diagnosis software into the robot control system. Analog low-pass filtering is realized at a cut-off frequency $f_c = 40\,\text{Hz}$. Digital filtering to generate the derivative $\dot{\omega}$ is performed at $f_c = 20\,\text{Hz}$.

For parameter estimation the DSFI procedure (discrete square-root filtering in information form) is applied because of its good numerical properties. The forgetting factor λ is set to 0.99. Figure 8.6 shows the parameter estimates after starting the estimation procedure. They converge within one movement cycle to constant values.

Now process coefficient information is written into a symptom buffer at fixed time intervals or on request. Nominal values are supposed to be known from the prior training.

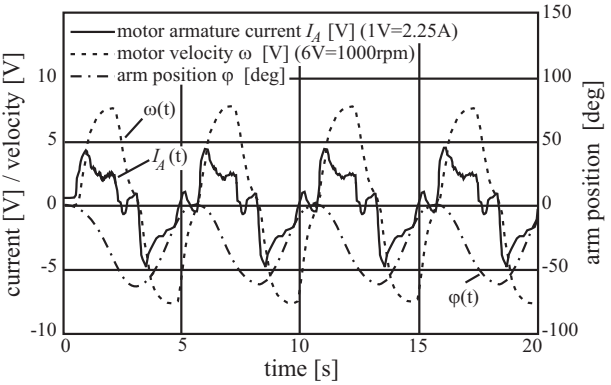


Fig. 8.5. Time history of measurements of basic axis 1 in the case of a periodic point-to-point movement

A membership function for positive and negative changes of the symptoms is introduced:

$$\mu_i(S_i) = \frac{1}{(b_i - a_i)}\,(S_i - a_i) \tag{8.4.1}$$

such that

$$\begin{aligned} \mu_i(S_i) = 0 \text{ for } S_i < a_i \quad &< \text{no significant increase} > \\ \mu_i(S_i) \in [0,1] \text{ for } a_i < S_i < b_i \quad &< \text{significant increase of size } \mu_i > \\ \mu_i(S_i) = 1 \text{ for } S_i \geq b_i \quad &< \text{significant increase} > \end{aligned}$$

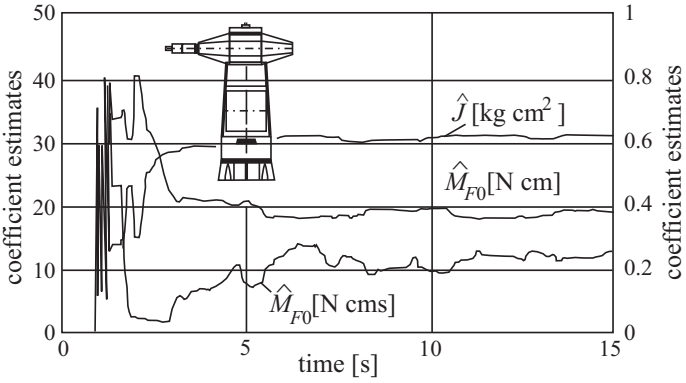


Fig. 8.6. Parameter estimates with the signals of Figure 8.5

Herewith a_i and b_i are lower and upper bounds (thresholds) which are selected based on statistics of the features (process parameters) like the standard deviations σ_{si}, e.g. $a_i = \kappa_1 \sigma_{si}$, with, e.g. $\kappa_1 = 2$ or 3 and $b_i = \kappa_2 \sigma_{si}$, with, e.g. $\kappa = 5$. Symptoms with different signs are treated as separate symptoms.

Now the introduction of different faults into the robot axis 1 is considered. After storing parameter set no. 60 fault F_1 was intentionally introduced by loosening the bracing screw between the mechanical components "spur gear" and "ring gear" as indicated in Figure 8.7. This fault was removed after recording parameter set no. 120. Sixty steps later fault F_3 was introduced by intentionally tightening the previously loosened bracing screw. Figure 8.8 shows the behavior of estimates of the friction coefficients and the corresponding membership functions.

Various investigations have shown that changes in the mechanical states of the different axes can be detected at an early stage by this model-based approach without adding new sensors. An overview of the obtained results is given in Table 8.1, indicating the coefficient deviations in the case of different faults, valid for each of the robot´s axes. This table shows that a certain amount of diagnosis can be directly performed by pattern recognition. Nevertheless, it will be illustrated how the heuristic approach will improve the depth and reliability of diagnosis.

8.4.2 Fault diagnosis with analytical and heuristic knowledge

To illustrate the use of the heuristic part of fault diagnosis it is assumed that the symptom buffers for the analytic symptoms indicates $\mu(S_{3a}) = 1$ and $\mu(S_{5a}) = 0.2$, i.e. a significant decrease of dry friction and a weak decrease of viscous friction which can be caused by a large overload or by a fault during last maintenance. According to the fault-symptom tree in Figure 8.4 one obtains with the additional membership functions

$$\mu(\bar{S}_{5a}) = 1 - 0.2 = 0.8; \quad \mu(S_{S4h}) = 0; \quad \mu(E_5) = 0$$

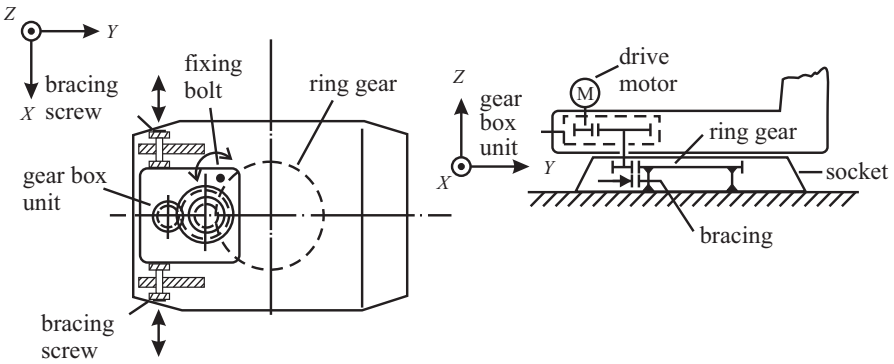


Fig. 8.7. Gearing scheme of axis 1, indicating fault location in the case of faults F_1 and F_2

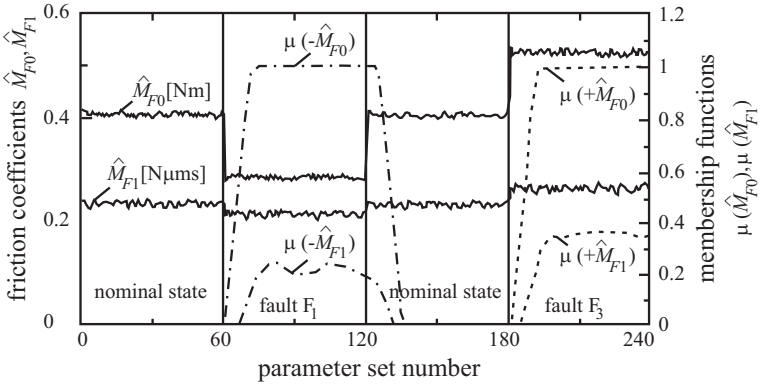


Fig. 8.8. Behavior of the friction coefficient estimates and the corresponding membership functions in the case of faults F_1 and F_3

Table 8.1. Process coefficient deviation caused by different mechanical faults in the drive chains of the investigated industrial robot. $+/++$ increase of coefficient: weak/strong; $-/--$ decrease of coefficient: weak/strong; 0 coefficient scarcely influenced

	θ_j			
F_j	$\Delta\hat{J}(E_0, m_L)$	$\Delta\hat{M}_{F0}$	$\Delta\hat{M}_{F1}$	$\Delta\hat{M}_{G0}$
F1	0	$--$	$-$	0
F2	0	$--$	$-$	0
F3	0	$++$	$+$	0
F4	0	$-$	$--$	0
F5	$++$	$+$	0	$++$
F6	$+$	$+$	$++$	0

by applying fuzzy logic with min operator for AND connections and max operator for OR connections

$$\mu(E_2) = 0.8; \quad \mu(E_1) = 0.8; \quad \mu(E_3) = 0$$

The last maintenance was just recently: $\mu(S_1) = 0.2$ or $\mu(\bar{S}_1) = 1-0.2 = 0.8$. Further $\mu(S_3) = 0.7$ (collision probably occurred) and $\mu(S_2) = 0.5$ (medium operating hours) leads to

$$\mu(F_1) = 0.7; \quad \mu(F_2) = 0.2; \quad \mu(F_3) = 0$$

The accumulation of the fault membership functions then leads to the conclusion that the fault F_1 arose, a loosened bracing screw of the spur gearing.

This example shows that a fuzzy-logic diagnosis procedure allows one to integrate analytic and heuristic knowledge in a transparent way.

8.5 Conclusions

If the axes of industrial robots perform dynamic motions, parameter estimation of a second-order differential equation leads to several features which can be compared to the normal behavior for a certain motion and effector load. Based on the determined symptoms several faults in the mechanics can be diagnosed. Combining this analytical knowledge with heuristic knowledge from maintenance personnel via membership functions allows one to confirm the analytical obtained results and to expand the diagnosis procedure.

9

Fault diagnosis of machine tools

The efficiency of manufacturing systems depends to a high degree on the reliability and availability of metal-cutting machine tools. Therefore, the detection and diagnosis of incipient and abrupt faults is of high importance. Statistics with regard to failure causes show for computer-numerical-control (CNC) drilling machines that tool faults constitute 27%, CNC faults 16%, mechanical faults 5%, electrical faults 4% and others, like organizational faults 34% and lack of orders 14%, [9.10], [9.40]. Hence, tool wear, breakage and collision contribute considerably to machine tool failures. Failure statistics for turning machine tools and machinery centers around 1993 showed that CNC and electrical failures constitute about 8%, and failures in the mechanical parts, like tool carriers 25%, workpiece handling 16%. These numbers underline the importance of automatic supervision or condition monitoring of machine tools. It is not only the breakdown of the manufacturing process that counts, but also damaged workpieces and tools.

In the following it is shown how model-based concepts allow one to detect and diagnose faults by using mainly standard sensor equipment or by adding additional sensors which can easily be implemented. The case studies consider main drives, feed drives, drilling, milling and grinding.

9.1 Structures of machine tools

In the following, machine tools are considered which operate with shape transformation processes removing metal from a workpiece to produce a mechanical product. The metal-removing processes include operations such as turning, drilling, milling and grinding. Machine tools can be divided into conventional machine tools, production machine tools and CNC-machining centers, [9.39]. *Conventional machine tools* are designed to perform one or several operations. They are used for general purpose machining of small lots of parts and are typically found in machine shops. *Production machine tools* are developed for high-volume manufacturing systems to perform one or a sequence of operations repetitively. They are usually numerically controlled and often connected to automatic material handling systems. *CNC-machine centers*

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_9, © Springer-Verlag Berlin Heidelberg 2011

are highly automated machine tools that produce a variety of more complex parts. They have different cutting tools e.g. in a tool holder and are capable to change cutting tools and to operate with simultaneous motions of different axes. Because of their flexibility they are economical for medium production volumes.

The basic construction of a machine tool consists of the stationary and moving mechanical structure, the drives and automatic control. The *mechanical structure* is typically based on stationary bodies, like bed, column, bridge and gear box housing, and on moving bodies, such as tables, slides and guideways, spindles, gears, bearings and carriages. These parts have to possess high mechanical rigidity, thermal stability and good vibration damping, in order to minimize static and dynamic deformations.

The *drives* consist of spindles and feed drives, see Figure 9.1. A spindle drive provides the required torque and speed to a rotating spindle shaft carrying the cutting tool with a clamping device. Box spindles are driven externally by an electric motor, belt drives, clutch and gears. If the electric motor is integrated into the spindle housing as a direct drive this is called a motorized spindle.

The relative motion between the tool and the workpiece is achieved by moving the table or carriage, spindle or column separately or together by *feed* or *axis drives*. These feed drives consist of a drive motor and a mechanical transmission. In the case of a driven table, this table is connected generally to a nut and a lead screw. The screw is driven either directly or via a gear system or belt drive.

The motor can be electric, hydraulic or pneumatic. Mostly AC or DC rotary motors are used or electric linear motors.

Automated machine tools require an *automatic CNC control unit* where according to software programs the whole machine tool is digitally controlled. This includes, e.g. the speed control of the spindles with speed sensors, the position and speed control of the feed drives, with position sensors coordinating all setpoints such that certain surface conditions and geometries are generated, and interpolation programs for contour generation, monitoring and supervision functions.

Machine tools for cutting tasks like drilling, turning, milling or grinding consist usually of a main drive (spindle) and one or several feed drives. In the case of *drilling* or *milling* the drill or cutter is turned by a main drive and the workpiece is held in a chuck on a table and moved with a feed drive. For *turning* the workpiece rotates by the main drive and the turning tool moves with usually two feed drives. *Grinding* machines are characterized by a main drive for the grinding wheel moved with a feed drive and either chucked or rotating workpieces.

For further details on machine tool structures and technologies see, e.g. [9.2], [9.5], [9.7], [9.39], [9.52], Vol. 1.

A signal flow of the main and feed drive of a drilling or milling machine tool is depicted in Figure 9.2. The speed of the feed-drive motor influences the force from the moving workpiece to the rotating tool, and vice versa, the speed of the main drive influences the force acting on the workpiece. Hence, changes in the cutting process or in the drive train (motor, gear, spindle) are reflected in changes of the electrical current and speed of the driving motors. These relations can be used for model-based fault-detection methods. Therefore, the following sections describe

mathematical models of the main drive and the feed drive, before special cutting machine tools are considered.

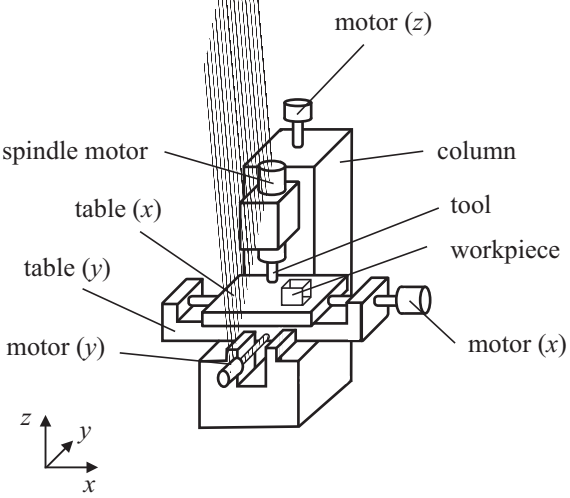


Fig. 9.1. Scheme of a machine tool for milling and drilling. One spindle drive and three feed drives

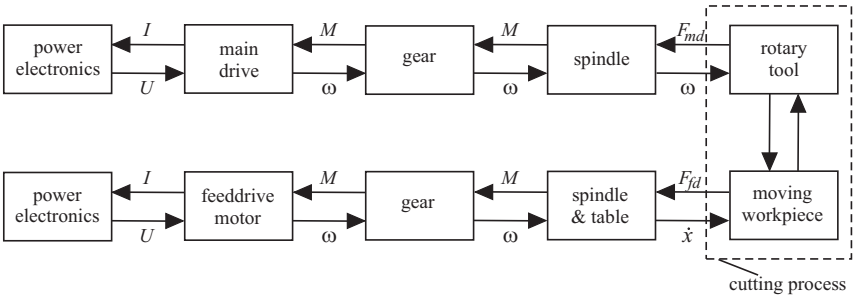


Fig. 9.2. Two-port representation of the signal flow of the main drive and one feed drive of a drilling or milling machine tool

9.2 Status of machine tools supervision

The automated workpiece processing through turning, drilling, milling and grinding requires that the cutting process operates according to the program of a numerical control (NC) unit and, after machining, satisfies the quality criteria. However, due to

the intensive interaction of the cutting tool and the workpiece deviations, faults and failures occur. Phenomena which influence the quality of the workpiece are chatter, tool wear, burr, heat generation and kind of chip formation and their transport. Examples of drastic events are tool breakage, workpiece slipping or collisions. Therefore, the continuous monitoring and fault diagnosis is of upmost importance for CNC machine tools.

A comprehensive survey of the technology and supervision of machine tools is given in [9.52], Vols 1 and 3. Condition monitoring and control as a special topic is treated in [9.48].

The cutting forces resulting from the contact between the tool and the workpiece depend on several variables like the depth of cut, the feed, the workpiece material and several empirical parameters, see Sections 9.4, 9.5 and 9.6. They directly inform on abnormal behavior. Therefore, the measurement of the cutting forces in one, two or three directions would allow a *direct fault-detection method*. However, they require force measurements at the workpiece clamping or the tool holding with, e.g. force cells based on strain gauges or piezoelectric stacks, are not only expensive but also weaken the stiffness. Though applied in laboratory machines they have not found application in practice. Then, *indirect detection methods* remain, based on easier applicable sensors, like main or feed drive current, accelerometers and sensors for vibration analysis, structure-borne noise, and acoustic emission.

Current measurement of the electrical motor drives is used as a standard way of limit checking to avoid overload of the motors and the machine and to protect against drastic failures in the workpiece/tool area.

There exist many publications on research results with these kinds of indirect fault detection, see, e.g. [9.52], Vol. 3 and a recent compilation in [9.47].

One general problem of these indirect methods is that they do not provide precise and robust information, also because they usually cannot be placed at the point of interest, [9.45]. Sometimes they allow a good indication of specific faults, but the coverage of other faults is not enough.

Examples of research in using ultrasound sensors (20–20,000 kHz) are [9.31]. The acoustic emission sensor is attached to the tool of the workpiece for precision-scale grinding.

The simultaneous use of several sensors, such as force, torque, vibration and spindle motor current for turning and drilling was reported in [9.8]. Extracted features from mean, variance and power spectrum were fused using principal component analysis and then treated in a fuzzy neural network, resulting in 80–95% success rate for the two cases investigated. A lot of research was performed on monitoring and fault diagnosis of *bearings*, and *rotating machinery*, see, e.g. [9.9], [9.11], [9.18], [9.24], [9.30], [9.33], [9.54], [9.55]. One example for rolling bearings in rotating machinery is the investigations in [9.30]. A comparison is given for position measurements of shaft eccentricity, acceleration and vibration sensors.

Usually *rolling bearings* produce vibrations due to varying compliance and imperfections of manufacturing and assembly, such as surface roughness, waviness, misaligned races and size-varying rolling elements. This holds even for new bearings and usually increases after the first defects start locally (cracks, pits, spalls).

The range of observed frequencies are for vibrations: 1 Hz to 25 kHz, ultrasonic noise: 20 kHz to 100 kHz, acoustic emission (surface waves): 100 kHz to 1 MHz plus shock pulses. For vibrations > 1 kHz acceleration measurement is best suited, e.g. by piezoelectric sensors. However, their placement is critical. Basic deterministic bearing frequencies can be calculated, depending on the location of the defect, geometry and speeds, but do not coincide with measured ones due to slippage and loading. In addition, interactions with other machine frequencies result in amplitude and frequency-modulated frequency bands, making analysis difficult. The analysis of the vibration signals is performed either in the time or frequency domain or both. A further problem is the influence of loading, because increasing load may decrease the vibration, and load is usually not known. Under ideal conditions prognostics of the expected lifetime seem to be possible after incipient bearing faults occur.

However, a general problem of bearing monitoring with vibration sensing is that frequently different bearings are assembled within rotating machinery, and also vibrations from other machine parts are generated (meshing gears, crankshafts, camshafts, chains, belts, etc.). Hence, a multitude of peaks in vibration spectra result, which are not easy to analyze and to trace back to their causes. If the resulting frequency peaks cannot be explained theoretically, there is still a chance to use vibration analysis experimentally, e.g. by introducing certain faults or failures artificially and training of feature extraction methods like neural nets. This is, for example used in testing combustion engines during assembling, [9.22].

Further research results for the fault detection of cutting processes are reported in [9.1], [9.4], [9.29], [9.38], [9.41], [9.43], [9.51].

A summary of research results and standard condition monitoring systems for *machine tools* is given in [9.52], Vol. 3. For turning and milling the cutting force shows only a small dependency on the wear. The passive cutting force seems to be more suitable. Therefore, the effective power calculated from the main drive is not sensitive enough for wear detection. Similar results are reported for drilling, but only with large drilling diameters. Structure-borne noise shows usually a broad and varying frequency spectrum. It is more effective for break detection than for tool wear. However, structure-borne noise is standard for grinding machines in detecting the grinding begin and supervising the trimming process.

This discussion shows that additional sensors for direct fault detection like force measurement are hardly used in practice, also because of high costs, changes of stiffness and wiring problems. Structure-borne noise may only be used in cases with a clear indication of faults. The following chapters are therefore dedicated to process-model-based indirect detection and diagnosis methods which use already implemented sensors or additional low-cost sensors.

9.3 Main drive

9.3.1 Two-mass model

As an example for a main drive the machine center of Figure 9.3 is considered (MAHO MC5). A speed-controlled DC motor drives a belt, a gear and tool spindle,

carrying a cutter or drill. Hence, a multi-mass–spring–damper system results with 6 masses. In the following, small deviations of the variables are considered such that linear models can be assumed. The dynamic behavior of the DC motor is modeled by

$$L_A \dot{I}_A(t) = -R_A I_A(t) - \Psi\omega_1(t) + U_A(t) \tag{9.3.1}$$

$$J_1 \dot{\omega}_1(t) = -\Psi I_A(t) - M_1(t) \tag{9.3.2}$$

with

L_A	armature inductance	U_A	armature voltage
R_A	armature resistance	I_A	armature current
Ψ	magnetic flux linkage	$\omega_1 = \dot{\varphi}$	motor speed
J_1	moment of inertia	M_1	load torque

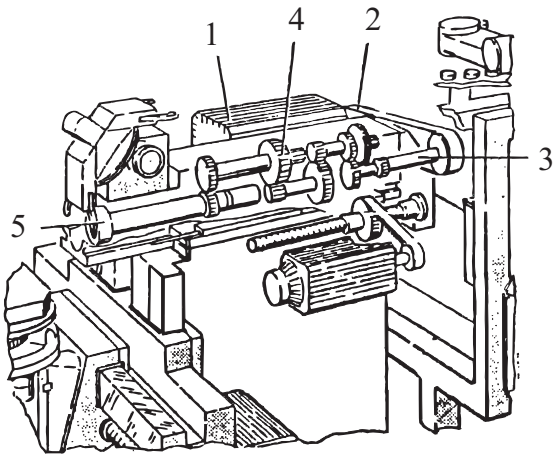


Fig. 9.3. Main drive of a machine center (MAHO MC5). 1 DC motor, 2 belt drive, 3 axle, 4 gear, 5 tool spindle

An analysis of the eigenfrequencies of the main drive shows, [9.13], [9.50], that the motor is able to excite frequencies in open loop $f < 80\,\text{Hz}$ and in closed loop $f < 300\,\text{Hz}$. The eigenfrequency of the belt drive is 123 Hz and those of shaft, gear and spindle 706, 412 and 1335 Hz. Hence, the dynamic behavior of the main drive is dominated by the motor and the belt drive and can therefore be modeled by a two-mass system with moments of inertia J_1 (motor plus belt-driving pulley) and J_2 (belt driven pulley, shaft, gear, spindle). The mechanical part of the main drive is then described by a linear state-space model

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) + \mathbf{F}\mathbf{z}(t) \tag{9.3.3}$$

with

$$\mathbf{x}^T(t) = [I_A(t), \varphi_1(t), \dot{\varphi}_1(t), \varphi_5(t), \dot{\varphi}_5(t)] \tag{9.3.4}$$

$$u(t) = U_A(t) \tag{9.3.5}$$

$$\mathbf{z}^T(t) = [M_6(t), M_F(t)] \tag{9.3.6}$$

M_6 : load torque; M_F : Coulomb-friction torque.

9.3.2 Parameter estimation

The parameters of the main drive can of course be determined from construction data. However, if not all parameters can be determined or if fault detection in normal operation, the estimation of the parameters from measured signals is desired.

To estimate the parameters of the main drive in idle running ($M_6 = 0$) based on measurements of accessible signals $U_A(t)$, $I_A(t)$, $\omega_1(t)$ and spindle speed $\omega_5(t)$ the following equations are used:

$$\begin{aligned} U_A(t) &= \Theta_1\omega_1(t) + \Theta_2 I_A(t) + \Theta_3 \dot{I}_A(t) \\ \Theta_1 I_A(t) - M_F(t) &= \Theta_4\dot{\omega}_1(t) + \Theta_5\dot{\omega}_5(t) \\ \omega_5(t) &= \Theta_6\dot{\omega}_1(t) + \Theta_7\omega_1(t) - \Theta_8\dot{\omega}_5(t) - \Theta_9\ddot{\omega}_5(t) \end{aligned} \tag{9.3.7}$$

with

$$\begin{aligned} &\Theta_1 = \Psi \quad \Theta_2 = R \quad \Theta_3 = L_A \\ &\Theta_4 = J_1 \quad \Theta_5 = iJ_2 \quad \Theta_6 = di/c \\ &\Theta_7 = i \quad \Theta_8 = d/c \quad \Theta_9 = J_2 i^2/c \end{aligned} \tag{9.3.8}$$

The armature flux linkage is first estimated by the first equation in (9.3.7) (or known from the data sheet). Then all process coefficients can be determined:

$$\begin{aligned} i &= \Theta_7 \text{ (gear ratio)} \qquad c = \Theta_5\Theta_7/\Theta_9 \\ J_1 &= \Theta_4 \text{ (motor)} \qquad d = \Theta_5\Theta_7\Theta_8/\Theta_9 \\ J_2 &= \Theta_5/\Theta_7 \text{ (spindle)} \end{aligned} \tag{9.3.9}$$

The derivatives of first and second-order for continuous-time parameter estimation were determined with state variable filters designed as Butterworth filters of 6th order with corner frequencies of 79.6 Hz and 47.8 Hz. The solution of the incremental rotation sensors was increased to 4096 slots for the spindle and 1024 slots for the motor. Sampling time was $T_0 = 0.5$ ms. The results with the parameter estimation method DSFI (discrete square-root filtering information form, see [9.13], [9.18]) for step functions of the speed are shown in Figures 9.4 to 9.7.

The motor coefficients on the motor side Ψ, R_A and L_A converge very fast, within about 2 s, the mechanical coefficients J_1, J_2, M_F, c and d a bit slower within about 5 s. After about 15 s all eight process coefficients converge to steady-state values and agree relatively well with theoretical determined values, [9.49].

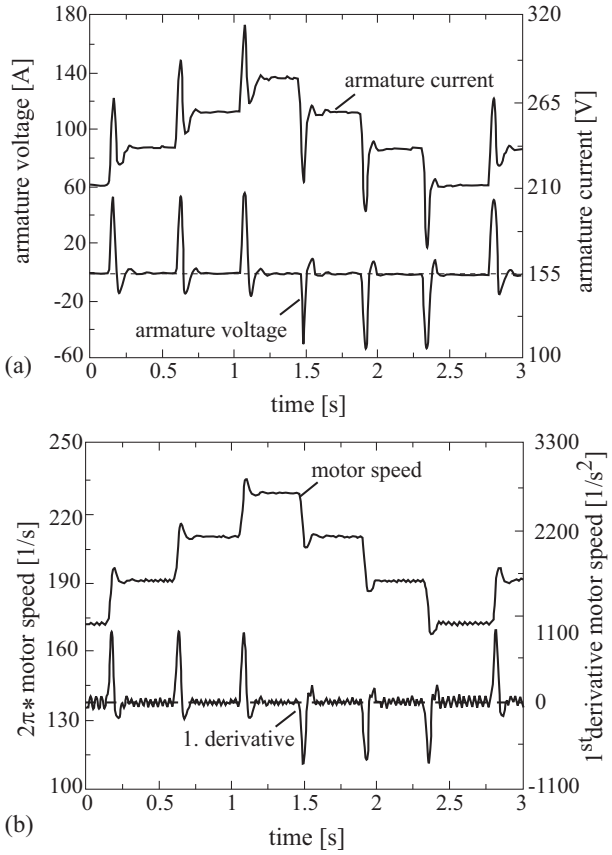


Fig. 9.4. Measured signals for the main drive. Step inputs for the setpoint of speed controller. (a) armature voltage and armature current; (b) motor speed ω and first derivative $\dot{\omega}$

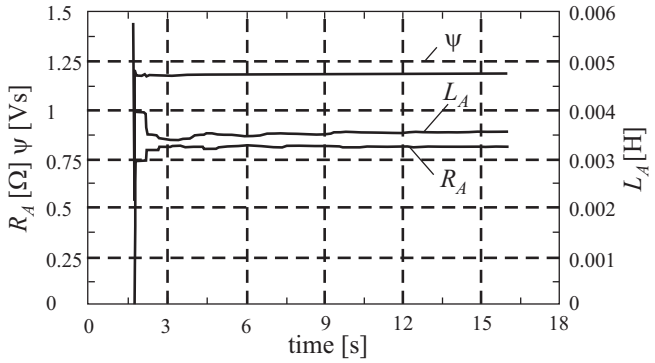


Fig. 9.5. Estimated process coefficients R_A, Ψ and L_A of the DC motor

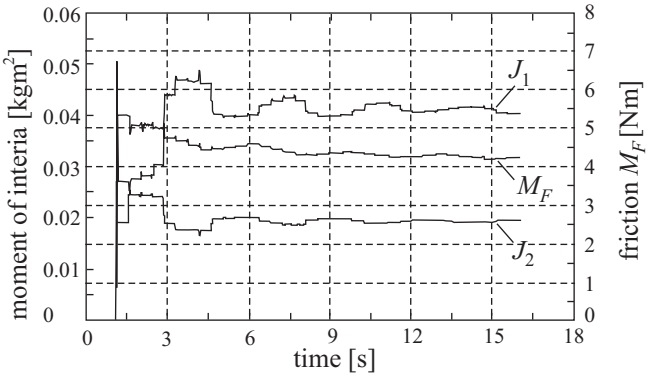


Fig. 9.6. Estimated process coefficients of the main drive. J_1 and J_2: moment of inertia of motor and spindle; M_F: dry-friction torque

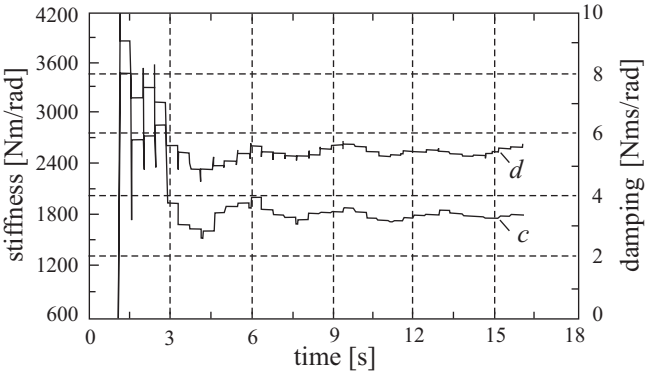


Fig. 9.7. Estimated stiffness c and damping d of the main drive

9.3.3 Fault detection with parameter estimation

To investigate the applicability of parameter estimation for the fault detection of the main spindle set-point changes with the speed control were made in idle running. Based on 60 data sets the standard deviations of the parameter estimates were between 0.01 and 10%. The resulting changes of the parameters are shown in Table 9.1 for the most relevant parameters J_2, M_F and c. F1 is not a real fault, but shows the sensitivity of the detection system and may serve to supervise the normal status. The belt defects are clearly detected and diagnosable, [9.49].

Table 9.1. Fault-symptom table for fault detection of the main drive with parameter estimation for speed setpoint changes in idle running. $+, ++$: small, large increase; $-, --$: small, large decrease; 0: no significant change

	J_2	M_F	c
F1 cutter $D = 120\,\text{mm}$	$++$	$+$	$-$
F2 increased belt tension	0	$+$	$++$
F3 belt with half width	0	$-$	$--$

9.4 Feed drives

9.4.1 Two- and three-mass model

For cutting and other machine tools, feed drives are used to move the machine table with high precision. The feed drives, usually controlled by a digital control unit yield the feed motion per turn for the cutting process. In the case of, e.g., milling the machine table carries the workpiece and in the case of turning, the cutting tool. For the design of precise position and trajectory control and for model-based fault detection accurate dynamic models of the feed drives are required. As an example the x-feed drive as shown in Figure 9.8 is considered, which moves the workpiece in a horizontal direction, see [9.28].

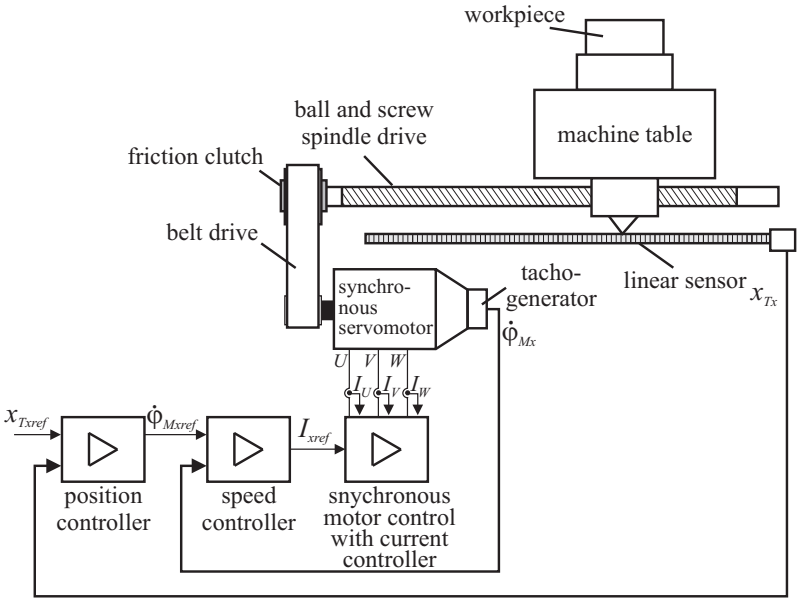


Fig. 9.8. Scheme of the x-feed drive control

The servomotor of the feed drive is a synchronous motor with constant excitation. After making some simplifying assumptions and use of a PI current controller the electrically generated torque M_{e1x} can be described by a first-order lag

$$T_{1Mx}\dot{M}_{e1x}(t) + M_{e1x}(t) = \Psi_x I_{x\,ref}(t) \tag{9.4.1}$$

where T_{1Mx} is the closed-loop time constant, [9.17]. The mechanical part of the feed drive motor yields

$$J_{Mx}\ddot{\varphi}_{Mx}(t) = M_{e1x}(t) - M_{Mf}(t) \tag{9.4.2}$$

with

J_{Mx} moment of motor inertia
φ_{Mx} motor angle
M_{Mf} motor friction

The mechanical part of the feed drive can be described as a system of coupled mass–spring systems. The feed drive consists of a motor shaft connected to a belt drive. The belt drive moves the feed screw by a friction clutch. The feed screw converts the rotational motion into a translational one and moves the machine table sitting on the feed screw nut. It consists of the belt drive with the rotational mass J_{Mx} of the motor with first pulley, the rotational mass of the second pulley and ball-and-screw spindle J_{Gx} and the translational mass m_{Tx} of the spindle drive screw and the machine table. Then a three-mass oscillator results as shown in Figure 9.9a). It is now assumed that the moment of inertia J_{Gx} of the pulley and spindle can be neglected because of the dominating masses of the motor and the table multiplied by the belt gear ratio i_x. This leads to the two-mass oscillator system shown in Figure 9.9b).

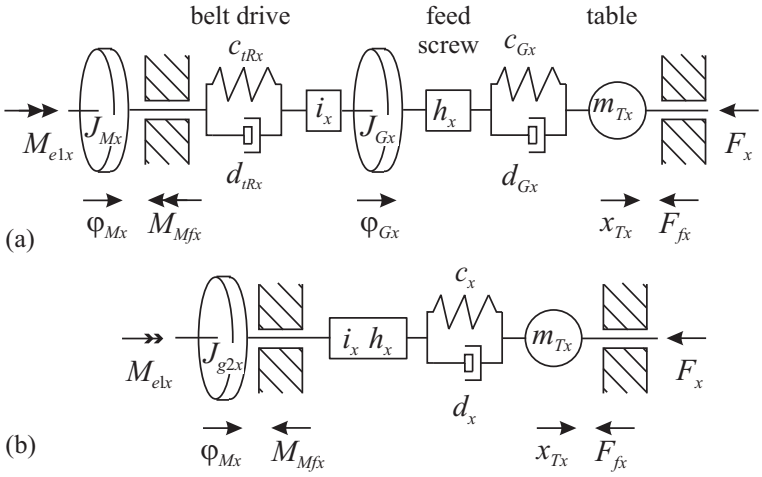


Fig. 9.9. Model structure of the mechanical part of the feed drive: (a) three-mass system; (b) two-mass system

Applying Newton´s second law or stating the force balance equation for the spindle, feed screw and table leads to

$$m_{Tx}\ddot{x}_{Tx}(t) = c_x\,(i_x h_x \varphi_{Mx}(t) - x_{Tx}(t)) + d_x\,(i_x h_x \dot{\varphi}_{Mx}(t) - \dot{x}_{Tx}(t)) - F_{fx}(t) - F_x(t) \tag{9.4.3}$$

with

m_{Tx}	table mass
x_{Tx}	table position
φ_{Tx}	angular position of ball screw drive (drive side)
h_x	lead of ball screw drive ($h = x_{Tx}/\varphi_{Tx}$)
c_x	overall stiffness (belt drive, spindle, screw)
d_x	overall damping coefficient
F_x	feed force in x-direction
F_{fx}	friction at slideways and ball screw
$i_x = \varphi_{Gx}/\varphi_{Mx}$	gear ratio of belt drive

The friction force is modeled with viscous and dry friction:

$$F_{fx}(t) = f_v \dot{x}(t) + f_c \operatorname{sign} \dot{x}(t) \tag{9.4.4}$$

The measured variables are the speed $\dot{\varphi}_{Mx}$ of the servomotor and the position x_{Tx} of the machine table by a linear incremental sensor with high resolution (500 slots/turn). To achieve high accuracy of the table position and fast dynamics the feed drive is controlled by a cascade control system with the motor speed controller as slave and the position controller as master, see Figure 9.10.

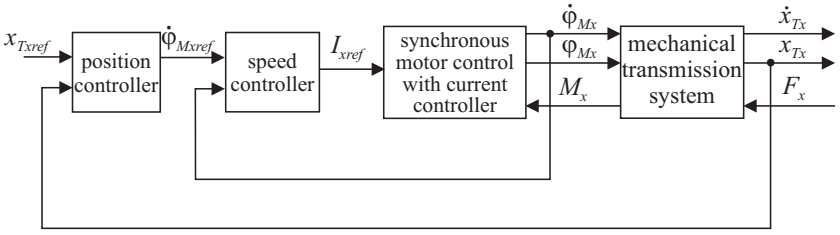


Fig. 9.10. Cascade control system for the position control

The reference value of the position x_{Txref} is calculated in the numerical control (NC) unit. The position controller is programmed as a proportional-acting controller:

$$G_{Px}(s) = \frac{\dot{\varphi}_{Mxref}(s)}{x_{Txref}(s) - x_{Tx}(s)} = K_{Px} \tag{9.4.5}$$

The underlying speed controller is realized analogously as a PI-controller with time lag

$$G_{nx}(s) = \frac{I_{xref}(s)}{\dot{\varphi}_{Mxref}(s) - \dot{\varphi}_{Mx}(s)} = K_{nx}\left(1 + \frac{1}{T_{Inx}s}\right)\left(\frac{1}{1 + T_{1nx}s}\right) \tag{9.4.6}$$

Based on the model structure of Figure 9.9b) a state representation can be derived with the state vector of 7th order:

$$\mathbf{x}^T = [x_{1n}, x_{2n}, I_x, \varphi_{Mx}, \dot{\varphi}_{Mx}, x_{Tx}, \dot{x}_{Tx}] \tag{9.4.7}$$

x_{1n} and x_{2n} are state variables of the speed controller. For further details see [9.28].

9.4.2 Identification of a feed drive

As several parameters of the feed-drive system cannot be determined precisely from data sheets, they were estimated based on measurements of the frequency response:

$$G_{\dot{x}}(i\omega) = \frac{\dot{x}_{Tx}(i\omega)}{\dot{\varphi}_{Mxref}(i\omega)}$$

with opened position control, but closed speed control, assuming the gear ratios i_x and h_x as known. For the experiments sinusoidal changes of the reference input of the speed controller were applied, with an added linear drift in order to avoid the nonlinearities of the dry frictions. As the parameter estimation method for the three-mass model Figure 9.9a) a numerical optimization method for the model output error was used (simplex algorithm). Figure 9.11 shows a comparison of the frequency response of the directly measured model with that of the parameter-estimated model of 9th order, given the linear model structure with three masses. The agreement is relatively good for $f \leq 120\,\text{Hz}$. Figure 9.12 shows the directly measured friction characteristics of the servomotor and the machine table. Finally an overall model as shown in Figure 9.13 was obtained. It was successfully applied for the reconstruction of the dynamic cutting forces of a milling process by measuring the positions φ_{Gx} of the pulley and x_{Tx} of the table and used for model-based fault detection of the milling tool, as described in Section 9.5.

9.4.3 Fault detection of a feed drive test rig

In order to investigate the fault detection with parameter estimation and artificially produced hard faults a special test rig was built, [9.12], [9.44]. Figure 9.14 depicts a schematic. A DC motor with 1.8 kW power drives a thread rolling drive via a toothed belt. The table with mass $m = 150\,\text{kg}$ is carried on slideways, where the friction can be changed by straining screws. The belt tension can be changed through tension screws by changing the distance between the shafts of the belt pulleys. Measured variables are the armature voltage U_A, the armature current I_A and speed ω_1 of the DC motor in closed speed control configuration. A one-mass model for the mechanical part turned out to be sufficient to estimate relevant process parameters. The models used are

$$\Psi_A \omega_1(t) = -L_A \dot{I}_A(t) - R_A I_A(t) + U_A(t) \tag{9.4.8}$$

$$\Psi_A I_A(t) = J\dot{\omega}_1(t) + M_{F1}\omega_1(t) + M_{F0}\,\text{sign}\,\omega_1(t) \tag{9.4.9}$$

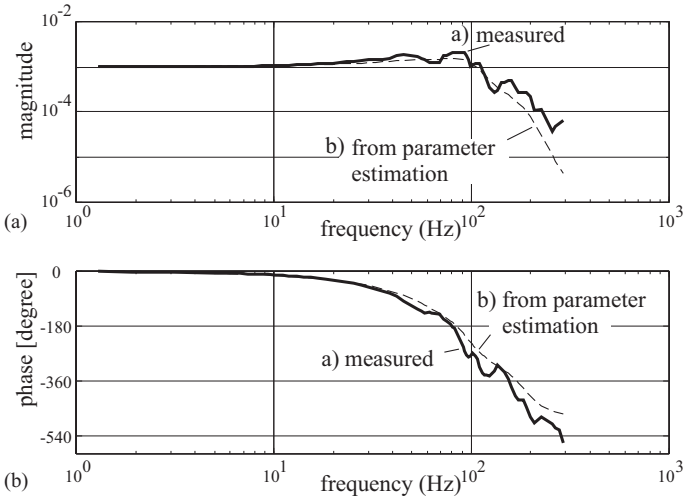


Fig. 9.11. Directly measured frequency response of the three-mass model with estimated parameters of the feed drive with $\dot{\varphi}_{Mxref}$ as input and $\dot{x}_{Tx}$ as output: (a) directly measured; (b) with parameter estimation

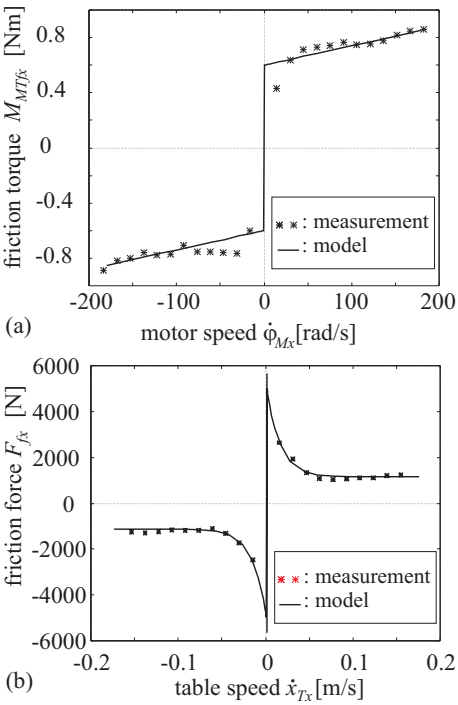


Fig. 9.12. (a) measured and identified friction characteristic of the servomotor; (b) measured and identified friction characteristic of the machine table

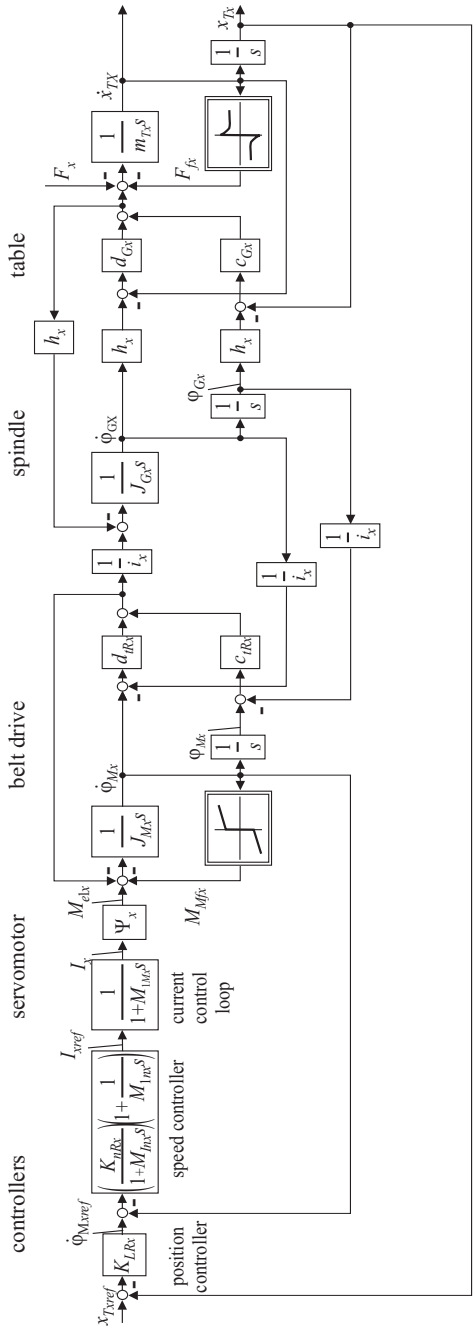


Fig. 9.13. Overall model of the feed drive of a machine tool with three masses (motor, spindle, table)

J is the overall moment of inertia including motor, spindle and table. Corresponding parameters for parameter estimation are

$$\begin{aligned}\boldsymbol{\Theta}_1^T &= [a_{11}, a_{10}, b_{10}] = [L_A/\Psi_A, R_A/\Psi_A, 1/\Psi_A] \\ \boldsymbol{\Theta}_2^T &= [a_{21}, a_{20}, a_{200}] = [J/\Psi_A, M_{F1}/\Psi_A, M_{F0}/\Psi_A]\end{aligned} \tag{9.4.10}$$

Based on the parameter estimates the following six process coefficients can be determined:

$$\begin{aligned} R_A &= a_{10}/b_{10} \quad J &&= a_{21}/b_{10} \\ L_A &= a_{11}/b_{10} \quad M_{F1} &&= a_{20}/b_{10} \\ \Psi_A &= 1/b_{10} \quad M_{F0} &&= a_{200}/b_{10} \end{aligned} \tag{9.4.11}$$

Figure 9.15 shows measured signals for a sinusoidal change of the reference value of the speed controller with amplitude 1000 rpm (peak-to-peak) and angular frequency $\omega_r = 3.1242$ rad/s after analog antialiasing filtering with a Butterworth filter of 8th order with corner frequency $f_c = 50$ Hz. Sampling time was $T_0 = 6$ ms. For continuous-time parameter estimation the 1st derivative $\dot{\omega}_m(t)$ was determined with a digital Butterworth state-variable filter of 4th order. The least-squares parameter estimates show a fast convergence for the motor parameters, Figure 9.16. Figure 9.17 depicts the motor parameter estimates during an increasing load through the straining screws. R_A and Ψ_A show significant changes, in dependence on the motor casing temperature, but less for L_A.

With increasing straining screw torques at the slideways the parameter estimates of the mechanical part show, that the dry-friction coefficient increases, but not the viscous-friction coefficient, Figure 9.18. However, both friction coefficients increase with increasing belt tension, Figure 9.19. Table 9.2 summarizes the changes of the parameter estimates for the faults. Hence, different patterns of deviations can be observed for the investigated faults and thus these faults can uniquely be diagnosed, [9.12].

Table 9.2. Fault-symptom table for different faults of the feed drive test rig. $+$, $++$: small, large increase; $-$, $--$: small, large decrease; 0: no significant change

parameter estimates	R_A	L_A	Ψ_A	J	M_{F0}	M_{F1}
F1 heating of motor	$++$	$+$	$--$	0	0	0
F2 defective commutator	$+$	0	$-$	0	0	0
F3 missing lubrication slideways	0	0	0	0	$++$	$-$
F4 large straining at slideways	0	0	0	0	$++$	0
F5 too large belt tension	0	0	0	$-$	$++$	$++$
F6 defective belt	0	0	0	0	$--$	$++$
F7 overload on table	$+$	0	$-$	$+$	$++$	$+$

It is especially noted that the fault detection of both the electrical motor and the connected machinery is based on easily measurable signals of the motor only. By

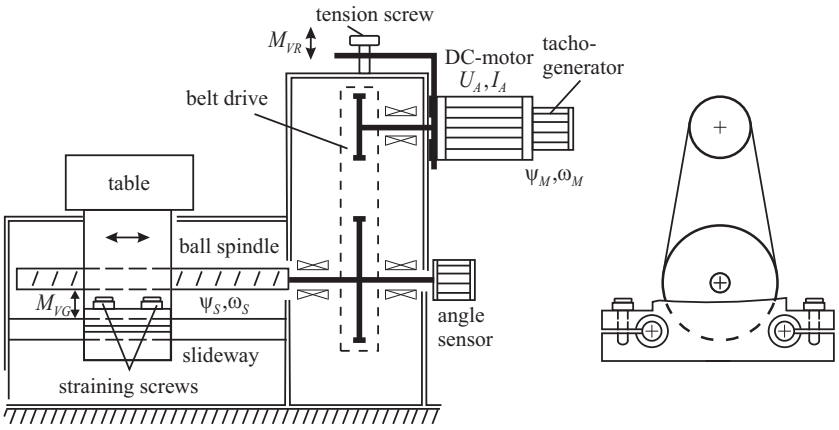


Fig. 9.14. Scheme of a feed drive test rig

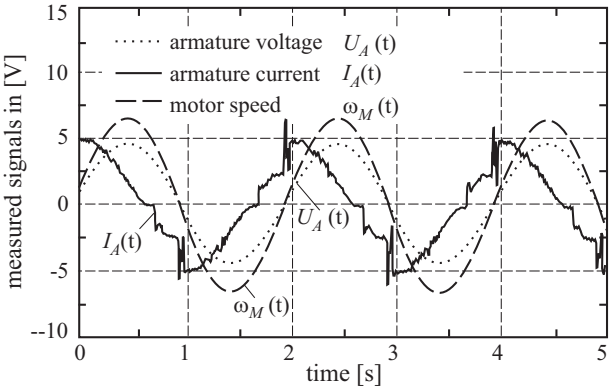


Fig. 9.15. Measured signals of the feed drive test rig. U_A: armature voltage, I_A: armature current, ω_m: motor speed (all signals in [V])

applying mathematical models of the motor and the mechanics, parameter-estimation methods allow one to use the "motor-as-sensor-principle," proposed in [9.14].

If the rotational positions φ_1 of the motor and φ_2 of the spindle pulley are measured additionally, then based on their difference $\Delta\varphi = \varphi_1 - \varphi_2$ the stiffness c and damping d of the belt can be estimated. This allows one to isolate the causes of belt drive faults like wrong pre-tension or belt defects better from other faults of the feed drive, [9.12],

Another possibility is to analyze the frequency spectrum of the speed of the driven pulley, because the transverse and longitudinal oscillations change with belt tension and belt defects, [9.32].

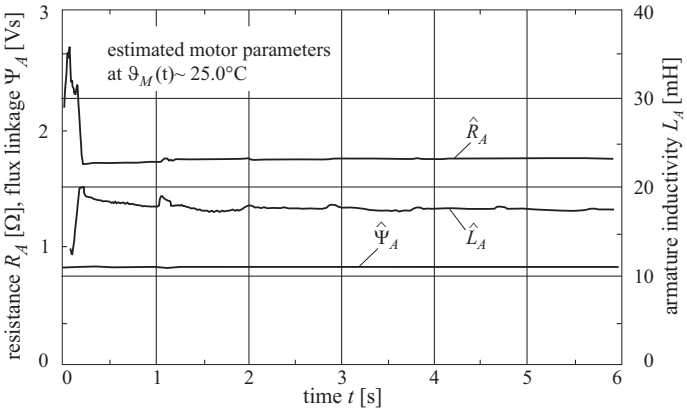


Fig. 9.16. Parameter estimates of the DC motor

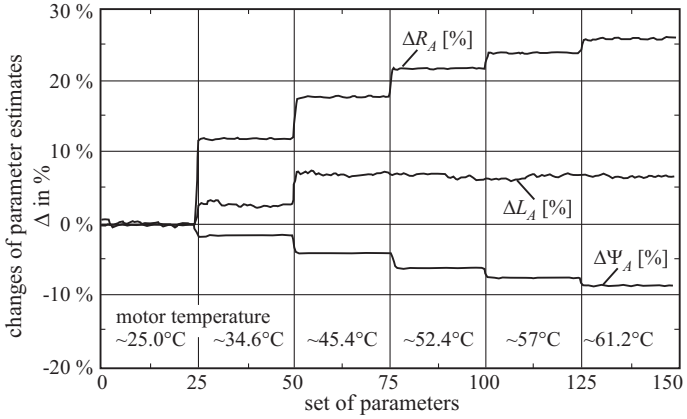


Fig. 9.17. Parameter estimates under the influence of motor temperature

9.5 Drilling machines

Drilling machines exist in a large variety of types, e.g. as stand-alone machines, in machine tool centers or as multi drilling heads in transfer lines. In the following the process of drilling is considered for a standard drilling process.

9.5.1 Models of the drilling process

For the development of model-based fault detection of drilling the machining center of Figure 9.3 is used. The drill is held by the main drive axis and moved with its z-feed drive against the workpiece, see Figure 9.1.

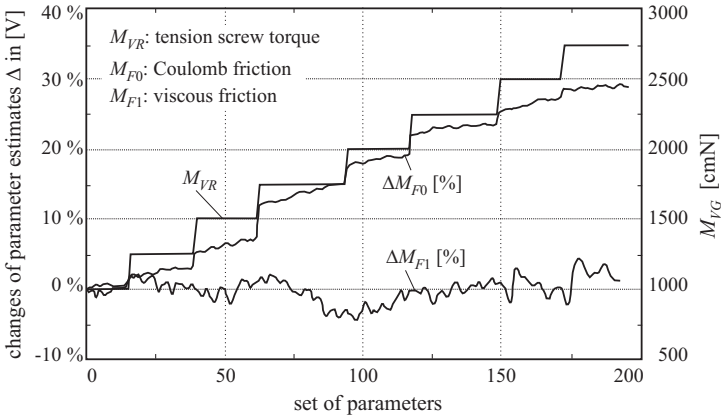


Fig. 9.18. Changes of the friction coefficients $\hat{M}_{F0}$ and $\hat{M}_{F1}$ during stepwise changes of the straining screws for the slideways. M_{VG}: straining screw torque

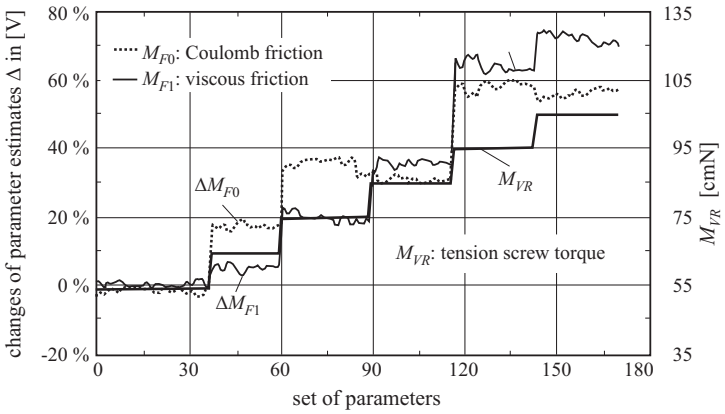


Fig. 9.19. Changes of parameter estimates during stepwise changes of the tension screws for the belt M_{VR}: tension screw torque

a) Static model

Drilling is a form of metal cutting with a circular motion. The tool works with a feed motion in direction of the rotation axis. The cutting forces on each edge can be decomposed into longitudinal and radial direction, see Figure 9.20. The longitudinal forces result in a feed force for which approximately holds, see [9.37]

$$F_f = A_f k_f \tag{9.5.1}$$

with A_f cutting face, k_f specific feed force.

The specific feed force k_f depends on material type, feed velocity and in particular on tool wear. These influences are taken into account by adding correcting factors to the basic specific feed force $k_{f1.1}$:

$$k_f = k_{f1.1} k_w k_r h^{-m} \tag{9.5.2}$$

with $k_{f1.1}$ specific feed force related to a basic face and

k_w influence of tool wear
h depth of cut
m influence of material type
k_r all other influences.

The factor k_f describes the influences of the material type, of the feed velocity and of the tool angle κ, see Figure 9.20. The cutting face A_f is

$$A_f = d_B h \tag{9.5.3}$$

with d_B tool diameter.

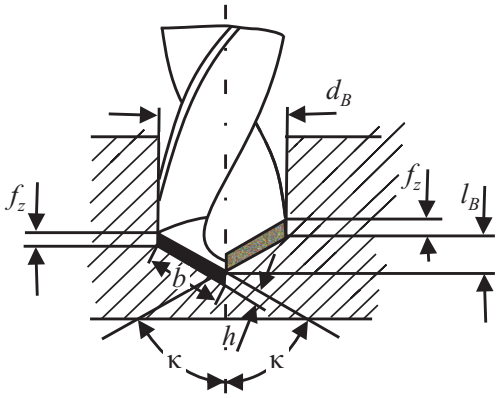


Fig. 9.20. Cutting geometry of the twist drill

Therefore the feed force F_f yields

$$F_f = k_{f1.1} k_w k_r d_B [f_z \sin\kappa]^{1-m} \tag{9.5.4}$$

with $f_z = \dot{f}\pi/\omega$ feed per tooth and

$\dot{f}$ feed velocity
ω angular velocity of the drill

and in a simplified form

$$F_f = \alpha \dot{f} \tag{9.5.5}$$

where α is a function of the feed velocity $\dot{f}$ and the influence of tool wear modeled by k_w.

Tool wear is caused by friction in the contact area between cutting edges and workpiece due to mechanical and thermal stresses. Tool wear is a collective description and depends on, [9.26]:

- edge fractures and deformations of the cutting edges
- adhesion
- diffusion
- oxidation.

Generally, the mentioned effects determine the behavior of friction as a whole and cannot be separated by force measurements.

b) Dynamic model

During initial drilling also a dynamic model can be derived between feed rate $\dot{f}$ as input and feed force F_f as output taking into account elasticity and damping of workpiece and drill, see [9.34],

$$T_{1f}\dot{F}_f(t) + F_f(t) = K_f\dot{f}(t) \tag{9.5.6}$$

with

$$\begin{aligned} K_f &= \alpha \\ T_{1f} &= \frac{\alpha}{c_f} + \tau_f \end{aligned} \tag{9.5.7}$$

where c_f is the stiffness of the workpiece–drill pair and $\tau_f \approx l_B/3\dot{f}$.

9.5.2 Fault detection of drilling

a) Wear detection

Experiments on a machining center with measured feed force F_f have shown that both parameters K_f and T_{1f} increase with increasing wear of the drill edges, [9.35], [9.51].

In order to save the feed force measurement the current I of the feed motors is used. For DC motors, universal motors and permanently excited synchronous motors it holds for the torque that

$$M_{mot} = \Psi I_{mot}$$

With the relation between the powers

$$P = M_{mot}\omega_{mot} = F_f\dot{f}\eta \tag{9.5.8}$$

where η is an efficiency coefficient and $\nu = \omega_{mot}/\dot{f}$ a gear ratio, the feed force can be replaced by

$$\begin{aligned} F_f &= k_m I_{mot} \\ k_m &= \Psi\nu/\eta \end{aligned} \tag{9.5.9}$$

leading to a first-order model

$$\frac{I_{mot}(s)}{\dot{f}(s)} = \frac{k_m}{1 + T_{1f}s} \tag{9.5.10}$$

Based on experiments with the machining center Figure 9.21 depicts the development of least-squares parameter estimates $\hat{k}_m$ and $\hat{T}_{1f}$ in dependence on the number of drills. Both parameters increase slightly until drill no. 212. The cutting edges of the drill were then artificially worn. Then a drastic increase especially of the gain $\hat{k}_m$ is observed. Hence, wear of drills can be detected by monitoring the current of the feed motor in dependence on the feed rate. Because the time constant T_{1f} is small, just low-pass filtering of the $I_{mot}(k)$ and $\dot{f}(k)$ and division to determine k_m is also sufficient for wear detection.

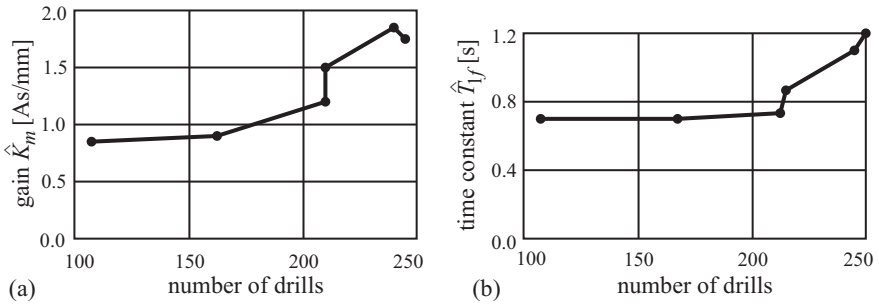


Fig. 9.21. Estimated gain $\hat{k}_m$ and time constant $\hat{T}_{1f}$ between current and feed rate in dependence on number of drills. Cutting speed $v_c = 18.5$ m/min; $f_z = 0.042$ mm; drill: $d_B = 5$ mm, HSS, St37; sampling frequency $f_0 = 5$ kHz

b) Breakage detection

Investigations on drill breakage have shown that the feed rate $\dot{f}$ changes earlier than the angular velocity ω_{mot} of the feed drive motor. Two residuals are calculated:

$$r_{\dot{f}}(k) = \dot{f}(k) - \hat{\dot{f}}(k) \tag{9.5.11}$$

$$r_{\omega}(k) = \omega_{mot}(k) - \hat{\omega}(k) \tag{9.5.12}$$

where $\hat{\dot{f}}$ and $\hat{\omega}$ are reconstructed by state observers with second-order models of the feed drive, with the current $I(t)$ of the feed drive as input. These observers can be considered as a *dedicated observer scheme* because different outputs drive the observers, which estimate the other process outputs, a principle for sensor fault detection, [9.6]. If then suddenly $r_{\dot{f}}$ passes a threshold earlier than r_{ω}, a drill breakage is indicated. Experiments with a 5 mm twist drill ($n = 1600$ rpm, $\dot{f} = 150$ mm/min), sampling rate $T_0 = 0.4$ ms have demonstrated that the detection time is about 15 ms, corresponding to 0.4 turns of the drill between breakage and detection, [9.16].

9.6 Milling machine

9.6.1 Models for the milling process

The considered machine tool is a machining center as shown in Figure 9.1 and Figure 9.3. The corresponding signal flow diagram is depicted in Figure 9.2 and the control system of the x-feed drive in Figure 9.8.

The most frequent faults in milling are wear and breakage of the inserts. Another fault is insert displacement which can, e.g., result from improper adjustment of the milling tool, see Figure 9.22.

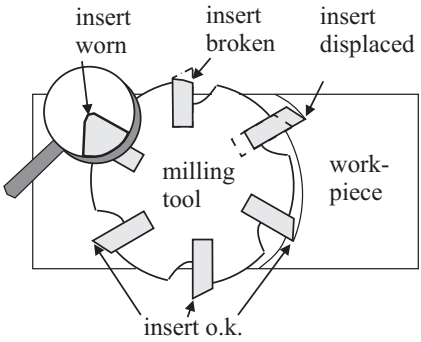


Fig. 9.22. Faults of a milling tool

Detection of faults in the milling process can be performed directly or indirectly. Direct methods usually imply great technical effort because special measurement equipment has to be applied to observe the cutter. Indirect methods are based on more-easily measurable signals which are influenced by faults. Because faults in milling have strong effects on the cutting forces most of the known indirect fault-detection methods use force sensors (dynamometers), see e.g. [9.3], [9.41], [9.42]. However, such force sensors are more suited to laboratory environment than to production machines because of their high costs and the mostly difficult mounting inside of the work cell. Besides direct force measurement also drive signals which are related to the forces can be used, see, e.g. [9.1], [9.15], [9.19], [9.38]. [9.1] and [9.38] show that the feed drive motor current can be used as remote force sensor to monitor milling respectively turning operations. But because of the damping of the drive train the use of current is only possible for lower-frequency parts of the force.

Faults in milling are usually detected if particular features generated from the measured signals cross predefined thresholds. Determination of these thresholds can lead to problems, especially if the cutting conditions are changed frequently. The reason is that most of the features are not only influenced by the faults, but also by changing operating parameters like spindle speed, feed per tooth, depth of cut, type of cutter, workpiece material etc. Hence, in order to avoid wrong decisions the thresholds have to be adapted if one of the cutting conditions is changed by the

operator. Further, most of the fault-detection methods have in common that they can only detect single types of faults, e.g., only wear or only breakage.

Therefore a model-based fault-detection method is described which is usable for distinct faults and independent of cutting conditions and direct force measurement. To obtain these properties, the proposed method uses analytical models of the feed drive and the milling process. The overall structure is given in Figure 9.23.

The next section describes the application of parameter estimation with three axis force measurements at the workpiece table to validate the models and to understand fault effects. Then the force measurement is replaced by only position measurements of the feed drive, following [9.27], [9.28]. Finally, fault detection with parity equations is applied.

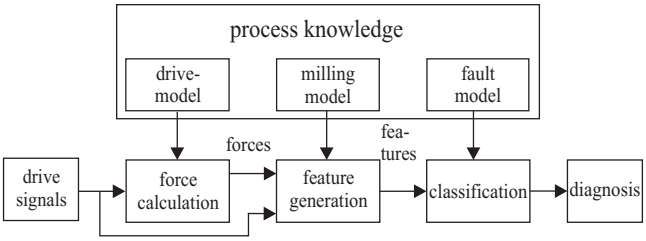


Fig. 9.23. Signal flow for the fault diagnosis of a milling machine

a) Feed drive model

During milling in the x-feed direction the feed force F_x causes an elastic deformation of the x-feed drive ball screw and an acceleration of the table. As a basis for the force calculation a simplified linear mass–damper–spring model of the ball screw drive is used, see Figure 9.24.

Taking into account a friction force in the guideway and spindle elasticity the force balance of the table leads to

$$F_{xc} = c\,(h\varphi_x - x) + d\,(h\dot{\varphi}_x - \dot{x}) - m\ddot{x} - F_f \qquad (9.6.1)$$

with

F_{xc}	[N]	calculated feed force (in x-direction)
φ_x	[rad]	angular position of ball screw drive (drive side)
x	[m]	table position
c	[N/m]	spring constant of ball screw
d	[Ns/m]	damping constant of ball screw
m	[kg]	table mass
F_f	[N]	guideway friction
h	[m]	lead of ball screw drive ($h = 0.01/2\pi$m)

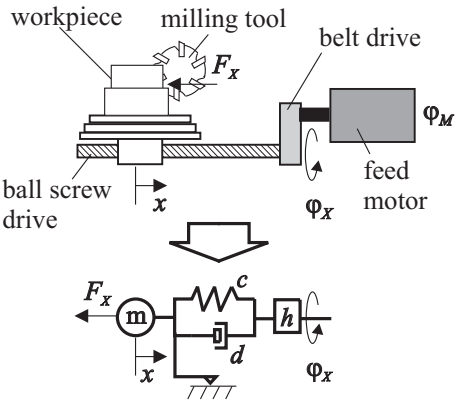


Fig. 9.24. Model of the ball screw spindle drive

To obtain the parameters c, d and m as well as the friction characteristics least-squares parameter estimation is used by minimizing the loss function V with N samples of the calculated force F_{xc} and the measured force F_x:

$$V = \sum_{k=1}^{N} (F_{xc}(k) - F_x(k))^2 \tag{9.6.2}$$

For identification of this model a direct force measurement is necessary. However, this procedure has to be performed only once, respectively in large time intervals to consider changes during the machine tool life cycle, e.g., aging.

To test the identification, measurements on the machining center were performed. For feed force measurement a Kistler force sensor (dynamometer) 9255A was used. The table position x could be obtained by a linear position encoder with 2000 pulses/mm, implemented for position control. To measure the angular screw drive position φ_x a Heidenhain encoder with 20,000 pulses/rev was mounted. Four milling operations with different cutting states were performed in order to excite the table:

M1: one insert broken
M2: one insert displaced
M3: all inserts totally worn
M4: all inserts new

The first three measurements were composed to one data set and used for identification. It turned out that exact modeling of friction F_f is crucial. The best results could be obtained by modeling F_f as the sum of a Coulomb term F_{fCoul} and linear term proportional to the table position F_{fpos}:

$$F_f = F_{fCoul} \operatorname{sign} \dot{x} + F_{fpos} x \tag{9.6.3}$$

The second term F_{fpos} is necessary to consider position-dependent friction effects which can, e.g., result from different states of wear on the guideway. The strong

influence of friction was already reported by [9.1] and [9.38] where current signals for force calculation were used.

To illustrate the performance of force calculation the data set used for identification and the forces calculated with the estimated parameters are given in Figure 9.25. Obviously, the correspondence is very good and the different cutting states can be well distinguished. Of course, the force calculation described above is valid only for the investigated feed drive.

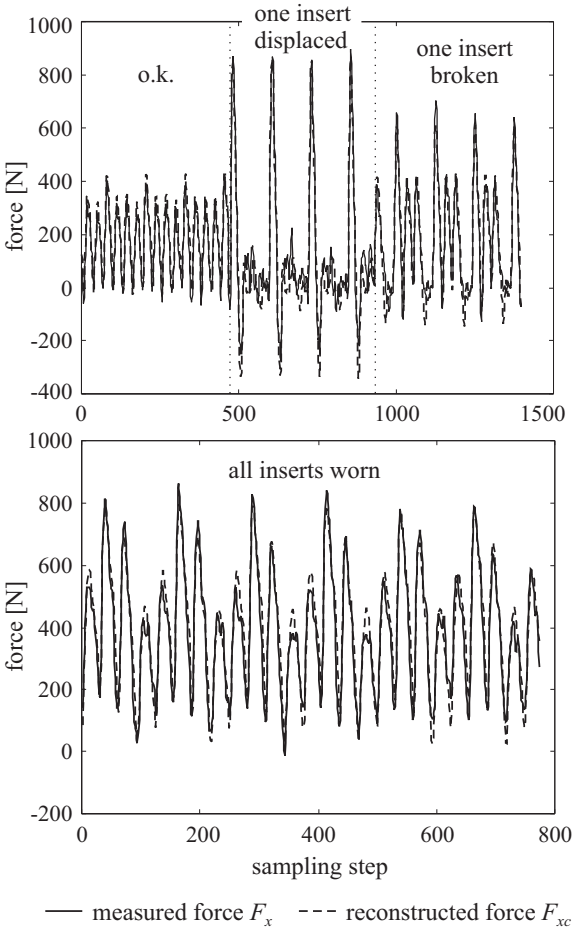


Fig. 9.25. Identification results for the feed drive

b) Cutting-force model

The tangential cutting force F_{ti} of an active cutting edge can according to [9.46] and [9.23], [9.25], be calculated by, compare Figure 9.26,

$$F_{ti} = k_t a_p h_i^{1-m_t} \tag{9.6.4}$$

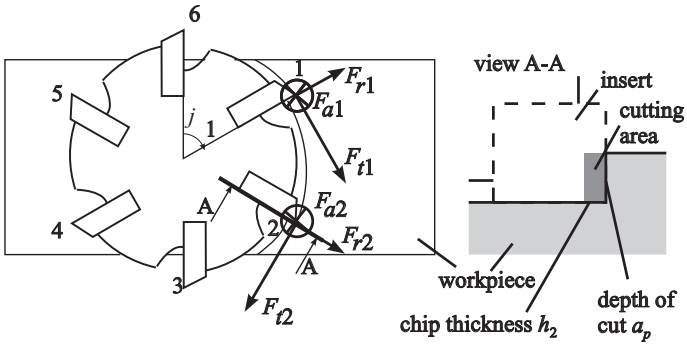


Fig. 9.26. Forces on a milling tool with orthogonal cutting inserts

For a milling cutter edge it yields for the cut thickness

$$h_i = f_z \sin \varphi_i \tag{9.6.5}$$

and therefore

$$F_{ti} = k_t a_p \, (f_z \sin \varphi_i)^{1-m_t} \tag{9.6.6}$$

with

k_t	[N/mm^2]	specific cutting force
a_p	[mm]	depth of cut
h_i	[mm]	local cut thickness
m_t		specific constant
f_z	[mm]	feed per tooth
φ_i	[deg]	angular position of tooth i
z		number of teeth (inserts)

The specific cutting force k_t depends on workpiece type, tool geometry, cutting speed and tool wear. These influences are taken into account with correcting factors

$$k_t = k_{t1.1} k_w k_r \tag{9.6.7}$$

with

$k_{t1.1}$	specific cutting force related to a basic face
k_w	influence of tool wear
h_r	other influences

The cutting radial force F_{ri} acting on tooth i at cutter rotation angle φ_i is

$$F_{ri} = \mu F_{ci} \tag{9.6.8}$$

In the fixed reference frame of the machine tool the cutting forces from one tooth i become

$$\begin{aligned} F_{xi} &= F_{ti}\cos\varphi_i + F_{ri}\sin\varphi_i \\ F_{yi} &= F_{ti}\sin\varphi_i - F_{ri}\cos\varphi_i \end{aligned} \tag{9.6.9}$$

The superposition of all cutting edges results in

$$F_t = \sum_{i=1}^{z} \delta(\varphi_i) F_{ti} \tag{9.6.10}$$

with

$$\delta(\varphi_i) = \begin{cases} 1 \text{ for } \varphi_1 \le \varphi_i \le \varphi_2 \\ 0 \text{ otherwise} \end{cases}$$

φ_1 entry angle of workpiece
φ_2 exit angle of workpiece

The resulting torque by the tool is

$$M_t = F_c r \tag{9.6.11}$$

with r effective cutter radius.

To evaluate the effect of faults the force equations for the faultless case are extended with correcting factors for each insert:

$$F_{tim}(k) = C_{ti} F_{ti}(k) = C_{ti} \left\{ a_P k_{t1.1} (f_z \operatorname{sign}(\varphi_i(k)))^{1-m_t} \right\} \tag{9.6.12}$$

$$F_{rim}(k) = C_{ri} F_{ri}(k) C_{ri} \left\{ a_P k_{r1.1} (f_z \operatorname{sign}(\varphi_i(k)))^{1-m_r} \right\} \tag{9.6.13}$$

$$F_{aim}(k) = C_{ai} F_{ai}(k) C_{ai} \left\{ a_P k_{a1.1} (f_z \operatorname{sign}(\varphi_i(k)))^{1-m_a} \right\} \tag{9.6.14}$$

with

C_{ti}, C_{ri}, C_{ai}		correcting quantities of the tangential, radial and axial force components
F_{ti}, F_{ri}, F_{ai}	[N]	tangential, radial and axial cutting-force components in faultless case
$F_{tim}, F_{rim}, F_{aim}$	[N]	modeled force components in tangential, radial and axial direction
$k_{t1.1}, k_{r1.1}, k_{a1.1}$	[N/mm^2]	specific force constants
m_t, m_r, m_a	[ ]	specific constants

Varying cutting conditions like depth of cut, feed per tooth and specific force parameters are included in the model and therefore do not influence parameters which model faults. The correcting factors depend therefore only on the process state, not on the cutting conditions. If the cutting process is free of faults, all correcting quantities are 1. Otherwise they differ from 1.

Finally the force in x-direction follows as

$$F_{xm} = \sum_{i=1}^{z} F_{tim}(k) \cos \varphi_i(k) + \sum_{i=1}^{z} F_{rim}(k) \sin \varphi_i(k) \tag{9.6.15}$$

where z is the number of inserts.

9.6.2 Fault detection of the cutter

a) Fault detection with force measurement

The insert forces $\mathbf{F}_{im}(k)$ are related to the measurable feed forces $\mathbf{F}_{Di}(k)$ by

$$\begin{bmatrix} F_{xm}(t) \\ F_{ym}(k) \\ F_{zm}(k) \end{bmatrix} = \begin{bmatrix} \cos\varphi_i(k) & \sin\varphi_i(k) & 0 \\ -\sin\varphi_i(k) & \cos\varphi_i(k) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{tim}(\varphi_i(k)) \\ F_{rim}(\varphi_i(k)) \\ F_{aim}(\varphi_i(k)) \end{bmatrix} \tag{9.6.16}$$

$$\mathbf{F}_{Di} = \quad \mathbf{T}(\varphi_i(k)) \quad \mathbf{F}_{im}(k)$$

where $\mathbf{T}$ is a transformation matrix depending on the angular position $\varphi_i(k)$ of the insert i.

For all inserts $i = 1 \ldots z$ it holds with (9.6.4) that

$$\mathbf{F}_{Di}(k) = [\mathbf{T}(\varphi_1(k))\mathbf{F}_{1m}(k) + \cdots + \mathbf{T}(\varphi_z(k))\mathbf{F}_{zm}(k)] \begin{bmatrix} C_{t1} \\ C_{r1} \\ C_{a1} \\ \vdots \\ C_{tz} \\ C_{rz} \\ C_{az} \end{bmatrix} \tag{9.6.17}$$

$$\mathbf{y}_m(k) = \boldsymbol{\psi}^T(k)\boldsymbol{\theta}$$

To estimate the unknown parameters $\boldsymbol{\theta}$ the output $\mathbf{y}_m(k)$ is replaced by the measured feed forces

$$\mathbf{y}(k) = \mathbf{F}(k) = \begin{bmatrix} F_x(k) \\ F_y(k) \\ F_z(k) \end{bmatrix} \tag{9.6.18}$$

and an equation error $e(k)$ is introduced, yielding

$$\mathbf{y}(k) = \boldsymbol{\psi}^T(k)\boldsymbol{\theta} + \mathbf{e}(k) \tag{9.6.19}$$

Now, $k = 1 \ldots N$ measurements are made ($N \geq 3z$) resulting in an equation system

$$\mathbf{y} = \boldsymbol{\Psi}(N)\boldsymbol{\theta} + \mathbf{e} \tag{9.6.20}$$

Minimizing the loss function

$$V = \mathbf{e}^T \mathbf{e} \tag{9.6.21}$$

leads to the least-squares estimate

$$\hat{\boldsymbol{\theta}} = \left[\boldsymbol{\Psi}^T \boldsymbol{\Psi} \right]^{-1} \boldsymbol{\Psi}^T \mathbf{y} \tag{9.6.22}$$

With this parameter estimation the correction factors C_t, C_r and C_a for each insert i are estimated by measurement of the feed drive forces $F_x(k)$, $F_y(k)$ and $F_z(k)$.

Table 9.3 shows experimental results with different faults. All faults could be detected and diagnosed, because the correcting factors increase for wear and insert displacement with different size and for breakage the factors become < 1, see Section c).

Table 9.3. Estimated correcting factors with direct measurement of feed drive forces

	correcting factors		
	C_{ti}	C_{ri}	C_{ai}
(1) fault free	1 ± 0.15	1 ± 0.15	1 ± 0.2
(2) wear	> 1.5	> 1.7	> 2
(3) breakage	≈ 0	≈ 0	≈ 0
(4) insert displacement	> 1.3	> 1.2	> 1.5

b) Fault detection with position measurements

The parameter estimation of the correction parameters C_i of (9.6.4) by measuring the table forces in all directions x, y and z has shown that the tangential and radial correcting parameters C_{ti} and C_{ri} can be estimated by using only the feed force F_x. In order to save an extra force measurement this feed force is now reconstructed by the feed drive model (9.6.1) and measurement of:

x table position
φ_x angular ball screw drive position on drive side
φ_s angular main spindle position (to know φ_i)

resulting in $F_{xc}(k)$. Then the insert forces $F_{tim}(k)$ and $F_{rim}(k)$ are calculated with (9.6.12), (9.6.13) and measurement of φ_s, but with the correction parameters as unknown s. (9.6.15) yields the milling model output $F_{xm}(k)$. Then a force error

$$e(k) = F_{xc}(k) - F_{xm}(k) \tag{9.6.23}$$

is used for least-squares parameter estimation of the correcting parameters $\hat{C}_{ti}$ and $\hat{C}_{ri}$ for each insert $i = 1, \ldots, z$, see Figure 9.27.

To test this procedure a large number of measurements was made on the machining center with a rectangular workpiece and the cutting conditions as in Table 9.4.

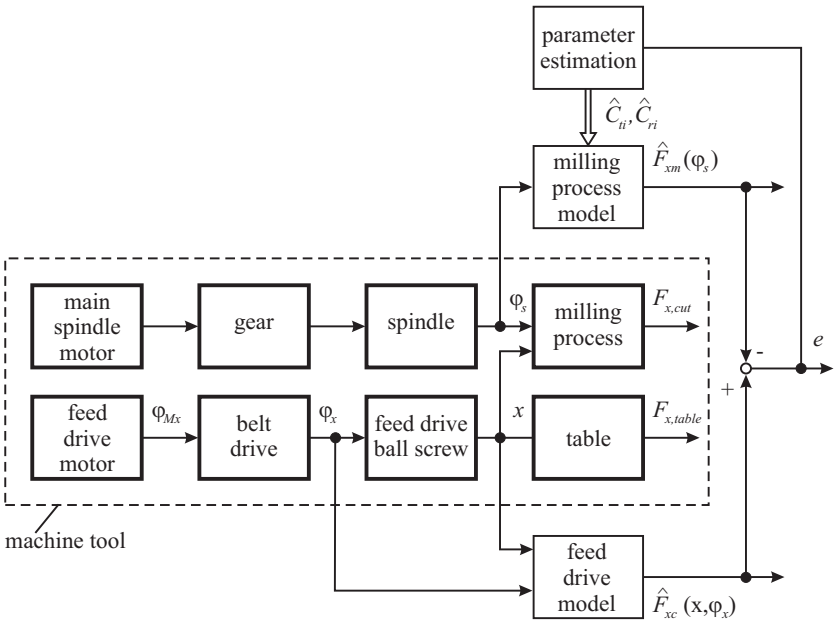


Fig. 9.27. Scheme for the parameter estimation of the milling forces through measurement of the feed drive angle φ_x on drive side and table position x

Table 9.4. Cutting conditions

workpiece	steel ST 37
cutter type	Widax M20
number of inserts	4
type of inserts	Widia TTM
tool radius	25 mm
milling type	end milling
depth of cut	1 mm
feed per tooth	0.04 – 0.4 rpm
spindle speed	625 – 1625 rpm

Feed per tooth and spindle speed were varied in the given range. During the measurement the faults were artificially generated. Even combinations of different faults were applied to the same milling tool, e.g. breakage of one insert and wear on the remaining inserts. Figure 9.28 depicts calculated and modeled forces calculated with the estimated correcting quantities. The effects of a breakage of insert 4 over three revolutions of the milling tool shows that a characteristic course can be seen. In the angular area of the broken insert the force is very small. Because of the higher load the force effecting the following insert is too high.

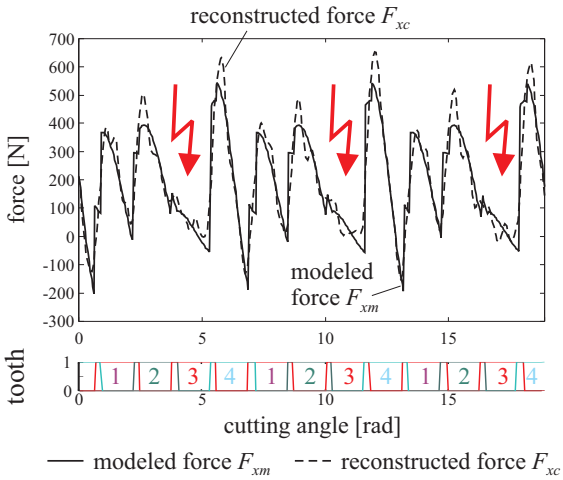


Fig. 9.28. Effect of breakage of insert 4 on the modeled feed force F_{xm} and the reconstructed feed force F_{xc}

The modeled feed force F_{xm} corresponds quite well with both the measured feed force F_x and the calculated feed force F_{xc}, see [9.28]. This can also be deduced from the similarity of the estimated correcting parameters.

Correcting parameters estimated with F_{xc}	Correcting parameters estimated with F_x
$C_{t1} = 1.58$	$C_{t1} = 1.58$
$C_{r1} = 1.31$	$C_{r1} = 1.49$
$C_{t2} = 0.93$	$C_{t2} = 1.08$
$C_{r2} = 1.31$	$C_{r2} = 1.27$
$C_{t3} = 0.97$	$C_{t3} = 1.17$
$C_{r3} = 0.86$	$C_{r3} = 0.93$
$C_{t4} = 0.43$	$C_{t4} = 0.06$
$C_{r4} = 0.03$	$C_{r4} = -0.08$

As expected, the correcting quantities of the inserts which are not influenced by the broken insert 4, i.e. inserts 2 and 3, are close to 1. The values of insert 4 are close to 0. Those of the next insert 1 are significantly higher than 1.

c) Fault diagnosis by classification

The resulting pattern of the estimated parameters can be used for fault detection. A classifier was developed with the task of distinguishing the following states:

(i) normal cutting
(ii) wear
(iii) breakage
(iv) radial insert displacement.

The classifier used consists of three parts, see Figure 9.29. In part 1 the correction parameter of the radial force C_{ri}, is evaluated. As the identification results have shown, C_{ri} remains comparatively uninfluenced in cases of breakage and insert displacement. Instead, C_{ri} is strongly dependent on tool wear. As long as wear does not cause breakages at the edges, there is no influence on the forces of the following insert. Therefore, wear can be detected by regarding only the corrective parameter of the particular insert. Due to the nature of wear, it is difficult to define a sharp threshold between "new" and "worn." Therefore, a fuzzy threshold is used to consider states between the two extremes.

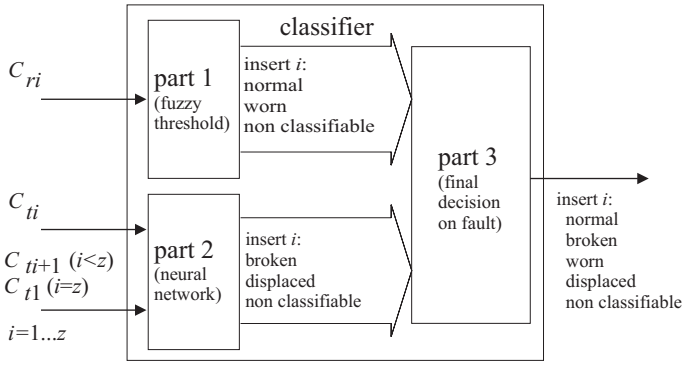


Fig. 9.29. Classification scheme for the estimated correcting parameters

If the previous insert is severely displaced or broken, C_{ri} can no longer be utilized to judge wear, and the output of part 1 is "non-classifiable."

In part 2 the correction parameter of the tangential force component C_{ti} is evaluated. As the identification results have shown, insert displacement and breakage lead to typical patterns of subsequent values of C_{ti}, see Figure 9.30.

In order to detect both faults, a neural network classifier is applied. Here, a multilayer perceptron (MLP) network with two inputs and two outputs is used, Figure 9.31. As inputs the tangential force correction parameter C_{ti} of the insert and of the following insert C_{ti+1}, are used. One of the binary outputs is activated in the case of insert displacement. The other output is activated in the case of insert breakage. In cases of no displacement *and* no breakage, the result in part 2 is "non-classifiable."

In part 3 the results of part 1 and part 2 are combined. A final diagnosis is produced by determining which of both results is dominant. If, for example, the output of the fuzzy threshold is "wear" and the output of the neural network is "non-classifiable" the final diagnosis in part 3 is "wear." The classification can at least diagnose one single fault and can give hints for two simultaneously appearing faults.

Extensive measurements with steel and aluminum as workpieces resulted in the classification results given in Table 9.5. Hence, the best overall results were obtained with the measurement of all three forces, good results with measurement of the feed

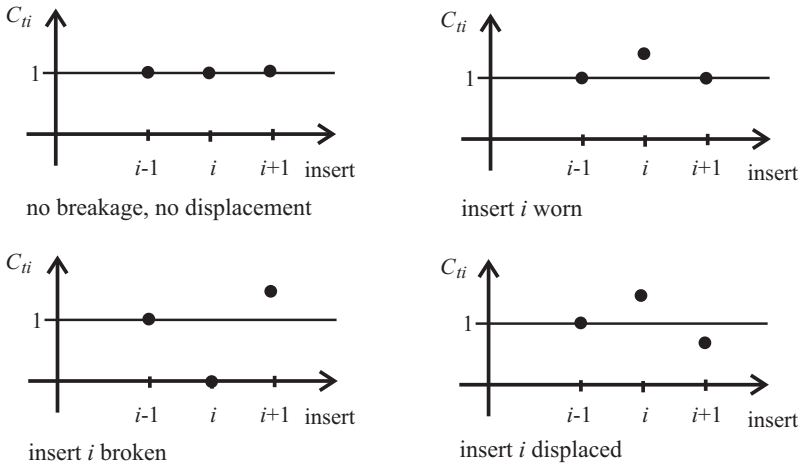


Fig. 9.30. Patterns of the tangential correction parameters C_{ti} for different faults

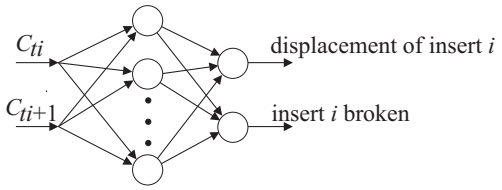


Fig. 9.31. Classification of tangential force correction parameters with a neural network to detect insert breakage and displacement

force and still good results with the model-based reconstructed feed force through feed drive position measurements.

Table 9.5. Diagnosis results for single inserts with parameter estimation of the correction parameters

milling tool status	percentage of correct classification [%]		
	direct measurement of F_x, F_y, F_z	direct measurement of F_x	reconstruction of force F_{xc}
(1) fault free	100	100	100
(2) wear	100	100	79
(3) breakage	89	100	89
(2) wear	100	67	67

d) Fault detection with parity equations

To result in a simpler fault detection as with parameter estimation the application of parity equations is considered. To save the expensive measurement of forces the angular ball screw position φ_x and the table position x are used to reconstruct the feed force $F_{xc}(k)$ with (9.6.1) and (9.6.3), which changes with faults. This force is compared with the expected force for normal operation by applying (9.6.15) leading to the force residual

$$r_F(k) = \Delta F_x(k) = F_{xc}(k) - F_{xm}(k) \tag{9.6.24}$$

The calculation of $F_{xm}(k)$ now contains no adaptation to the cutting conditions. The evaluation of the residuals is limited to a certain tool angle $\varphi_1 < \varphi_i(k) < \varphi_2$, where only one insert is in operation and to avoid large deviations during entrance and exit of the tool. As the maximal residuals are relatively large during normal operation (about 30% of maximal feed force) the thresholds have to be set to relatively large values, or adaptive thresholds, e.g. $|r_F| > \kappa |F_{xm}|$, $\kappa = 0.7$, have to be used.

Experiments have shown that this parity approach is able to detect broken inserts but not worn or displaced inserts.

The advantage of the parity equation approach is the smaller computational effort and the possibility of real-time application for small sampling time. The fault detection with parameter estimation allows a considerable larger fault detection and diagnosis performance. It is, however, computationally much more demanding. The described investigations were made with nonrecursive least-squares estimation and with a sampling time $T_0 = 0.4$ ms and a calculation time of 6.3 s. Therefore, it was not realtime, but timely enough for required actions after a fault was detected.

Additional investigations were made with *structure-borne noise sensors* mounted on the workpiece holder. The structure-borne noise was band-pass-filtered in ranges of 0–8 Hz, 8–16 Hz, 24–34 Hz and 34–50 Hz. The noise generated by the four inserts shows already in normal state much more deviation between each feed force measurement. Broken or displaced inserts could be detected in middle-frequency bands. However, wear could only be detected for small wear, not for larger wear. Hence, broken or displaced inserts are distinguishable, but displaced and worn inserts cannot be separated. As the structure-borne noise depends much on the individual cutting conditions and the wear status, and in addition the measurement location, it seems not to be a reliable source for fault detection of the milling cutting process. Much more details and comparisons are given in [9.27].

9.7 Grinding machines

As grinding of workpieces is one of the last processes in high-precision manufacturing of mechanical products the supervision and fault diagnosis of grinding machines is also advantageous for quality control. For example, tool wear and chatter oscillations have to be supervised. In the following a grinding machine for cylindrical surface grinding is considered as one example of different grinding processes, following [9.20] and [9.21].

9.7.1 Grinding-process models

Figure 9.32 shows a scheme of the considered grinding machine. The x-axis feed drive moves the rotating grinding wheel (peripheral speed $v_s \approx 45\,\text{m/s}$) with a feed motion $x_F(t)$ of the feed drive against the rotating workpiece (peripheral speed $v_w \approx 0.7\,\text{m/s}$). The resulting cutting normal force $F_N(t)$ causes a material removal $x_w(t)$. Three mass–spring–damper systems can be distinguished, the feed support system with the feed drive motor, the grinding wheel and the cylindrical workpiece with its suspension. For the grinding process the behavior within the contact zone is influenced by the stiffness c_x of the feed support, c_s of the grinding wheel and c_w of the workpiece. These stiffnesses are lumped together resulting in the overall contact stiffness

$$c_c = \frac{1}{1/c_x + 1/c_w + 1/c_s} \tag{9.7.1}$$

The normal cutting force $F_N(t)$ depends on the removal of workpiece material $x_w(t)$ and the grinding conditions.

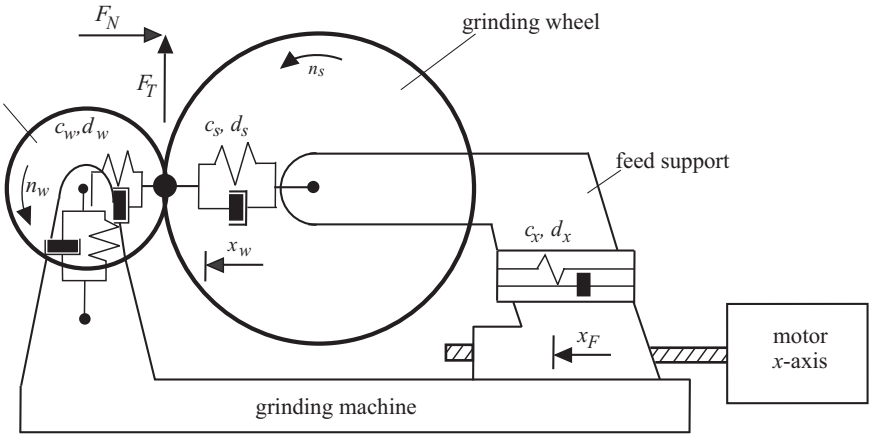


Fig. 9.32. Scheme of the cylindrical surface grinding machine

Based on [9.53] and [9.26] a simplified linear force equation can be assumed:

$$\begin{aligned} F_N(t) &= \frac{\alpha}{T_w} a_w(t) \\ a_w(t) &= x_w(t) - x_w(t - T_w) \end{aligned} \tag{9.7.2}$$

where $a_w(t)$ is the depth of a cut for one workpiece revolution T_w with $T_w = 1/n_w$, n_w the workpiece speed and α a grinding-force coefficient. According to [9.36] it holds for the cutting-force coefficient that

$$\alpha = bK \left(\frac{v_w}{v_s} \right)^{2\varepsilon_1 - 1} D^{1-\varepsilon_1} \tag{9.7.3}$$

with b width of the grinding zone, K a gain factor, D the grinding wheel diameter and $0.5 \leq \varepsilon_1 \leq 1$. The normal cutting force depends on the other side on the elastic deflection x_c in the contact area, which results from the feed position x_F, the removal x_w of the workpiece material and radial disturbances Δr_s of the grinding wheel and Δr_w of the workpiece surface. This leads to the force balance equation

$$\begin{aligned} F_N(t) &= c_c \left[x_F(t) - x_w(t) + \Delta r_s(t) + \Delta r_w(t) \right] \\ &= c_c x_c(t) \end{aligned} \tag{9.7.4}$$

Transforming (9.7.4) and (9.7.2) into the Laplace domain and eliminating $x_w(s)$ yields after back-transformation into the time domain

$$F_N(t) = a_1 F_N(t - T_w) + b_1 \left[x_F(t) - x_F(t - T_w) \right] \tag{9.7.5}$$

with

$$a_1 = \frac{\alpha}{\alpha + T_w c_c}; \quad b_1 = \frac{\alpha c_c}{\alpha + T_w c_c} \tag{9.7.6}$$

Hence, for a step input of the feed position the normal force $F_N(t)$ vanishes for $t \to \infty$ and in the case for a continuous depth increase $a_w(t)$ the normal cutting force remains constant. The transfer function following (9.7.5) is

$$G_{Fx} = \frac{F_N(s)}{x_F(s)} = \frac{b_1 \left[1 - e^{-T_w s} \right]}{1 - a_1 e^{-T_w s}} \tag{9.7.7}$$

The dead-time term can be replaced by the Padé approximation of first order:

$$e^{-T_w s} = \frac{1 - \frac{T_w}{2} s}{1 + \frac{T_w}{2} s} \tag{9.7.8}$$

This leads to

$$G_{Fx} = \frac{F_N(s)}{x_F(s)} \approx \frac{\frac{b_1}{1 - a_1} T_w s}{1 + \frac{T_w}{2} \frac{1 + a_1}{1 - a_1} s} \tag{9.7.9}$$

and with (9.7.6)

$$G_{Fx}(s) = \frac{F_N(s)}{x_F(s)} \approx \frac{K_D s}{1 + T_1 s} \tag{9.7.10}$$

$$K_D = \alpha \qquad T_1 = \left(\frac{\alpha}{c_c} + \frac{T_w}{2} \right) \tag{9.7.11}$$

Usually, the cutting force $F_N(t)$ cannot be measured. Therefore, the current $I_s(t)$ of the main spindle motor is used which can be assumed to be proportional to the torque, respectively to the grinding force, [9.20]:

$$F_N = k_I I_s \tag{9.7.12}$$

Then

$$G_{FI}(s) = \frac{I_s(s)}{x_F(s)} = \frac{\frac{\alpha}{k_I}s}{1 + T_1 s} \tag{9.7.13}$$

or in the time domain

$$I_s(t) = -T_1 \dot{I}_s(t) + K_s \dot{x}_F(t) \tag{9.7.14}$$

with $K_s = \alpha / k_I$.

9.7.2 Fault detection with parameter estimation

Experiments were performed with a Schaudt T3U grinding machine for a cylindrical workpiece. Figure 9.33 shows the measured signals during a three-phase grinding cycle with roughing, finishing and fine-finishing. The assumed process model is based on (9.7.14) with sampling time T_0:

$$y(k) = \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\Theta}} \tag{9.7.15}$$

the output signal

$$y(k) = I_s(k) \tag{9.7.16}$$

the measurement vector

$$\boldsymbol{\psi}^T(k) = \left[-\dot{I}_s(k)\ \dot{x}_F(t)\right] \tag{9.7.17}$$

and the parameter vector

$$\boldsymbol{\Theta}^T = \left[a_1'\ b_1'\right] = [T_1\ K_s] \tag{9.7.18}$$

By using least-squares parameter estimation and state-variable filters for determining the derivatives the gain $\hat{K}_s$ and time constant $\hat{T}_1$ can be estimated.

Based on these parameter estimates the cutting-force coefficient and the contact stiffness follows from (9.7.14) and (9.7.11):

$$\begin{aligned} \frac{\alpha}{k_I} &= \frac{\hat{K}_s}{T_w} \\ \frac{\hat{c}_c}{k_I} &= \frac{\hat{K}_s}{k_I\left(\hat{T}_1 - \frac{T_w}{2}\right)} \end{aligned} \tag{9.7.19}$$

Figure 9.34 presents the two parameter estimates for a grinding process with increasing number of workpieces. The cutting-force coefficient increases significantly and also somewhat weaker the contact stiffness. This is due to the wear of the grinding wheel because of blunt edges. In this case the grinding wheel should have been trimmed after the third working piece. These and other experiments have demonstrated that tool wear, wrong tools, missing cooling lubrication or trimming conditions can be detected with this parameter estimation approach, see [9.20].

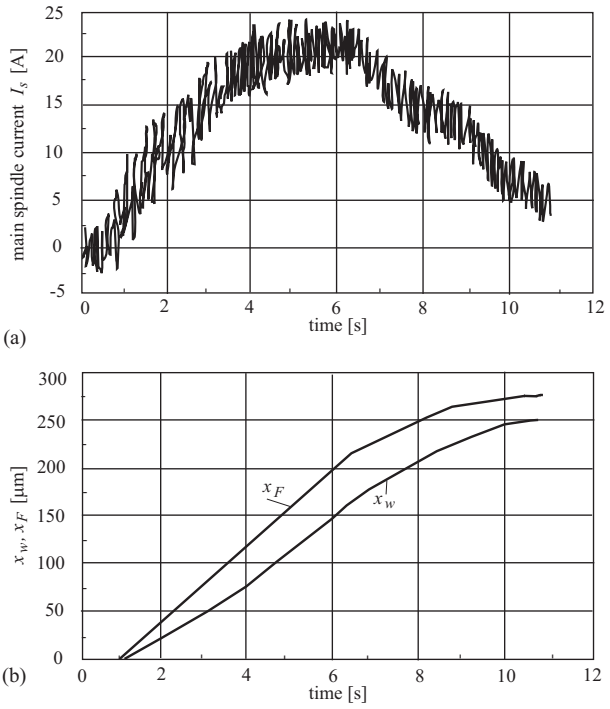


Fig. 9.33. Measured signals of the grinding machine: a) Main spindle current $I_s(t)$; b) feed position x_F, material removal x_w

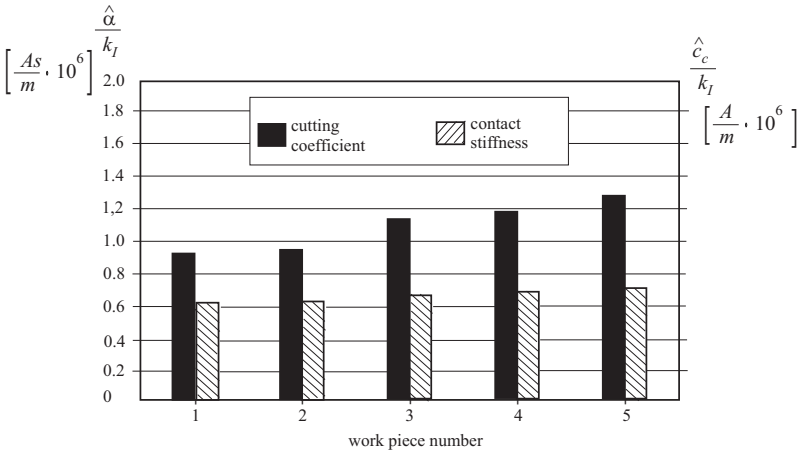


Fig. 9.34. Parameter estimates for grinding of workpieces. $\hat{\alpha}$ cutting coefficient, $\hat{c}_c$ contact stiffness, k_I motor constant

9.7.3 Fault detection with signal-analysis methods

To detect chattering of the grinding process a signal analysis of the grinding force or, if easier available, the main spindle motor current can be used. The signal analysis can be performed by FFT. However, to detect a limited number of frequencies a signal parameter estimation based on parametric signal models of ARMA type was applied in the form of the maximum entropy spectral estimation, described in [9.18], Chapter 8.1.6 and in [9.20]. Table 9.6 shows the estimated frequencies and amplitudes for a sampling rate of 250 Hz.

Table 9.6. Estimated frequencies and amplitudes of the main spindle current I_s with parameter ARMA signal models

frequency number	1	2	3	4	5	6	7	8
frequency f_i [Hz]	12.0	24.1	38.1	44.5	57.5	54.5	45.7	37.7
amplitude [A]	0.77	0.70	0.65	0.70	0.05	0.04	0.03	0.16
	belt drive main spindle		main spindle	main spindle motor				

Frequencies 1 and 2 are oscillations of the main spindle drive and the belt drive. f_3 is the rotating frequency of the main spindle and indicates chatter. f_4 is due to the main spindle motor which is related to the grinding wheel rotation by $i = n_M/n_s = 1.17$. The other frequencies do not contain important process information.

Figure 9.35 depicts the amplitudes of the main spindle frequency f_3 for increasing number of workpieces. The first three grinding cycles have too large an amplitude and are due to wrong trimming because of roundness errors of the grinding wheel and the workpiece. Then the process stabilizes to normal conditions (self-trimming effect) and increases for the last cycles with beginning of chattering.

Summarizing, a grinding process can be supervised with detailed indication of faults by model-based evaluation of feed position and main spindle current and frequency analysis of the main spindle current. With additional sensors, like for structure-borne noise, more information can be obtained [9.52].

9.8 Conclusions

The described results of several research projects for drilling, milling and grinding machine tools have demonstrated that model-based fault-detection methods in the form of parity equations and parameter estimation allow the fault diagnosis of cutting processes and machine tool parts. Especially, the feed drives and their position and current signal enable one to estimate forces of the cutting process or of the friction of movable parts, like slideways and ball screws. This holds also for the signals of the main drives. Hence, the application of a process-model-based “drive-as-sensor

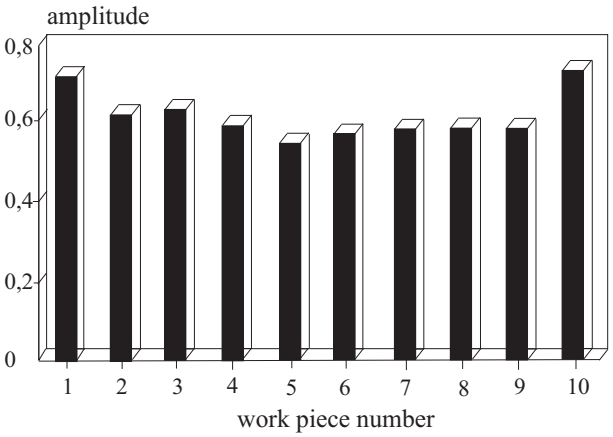


Fig. 9.35. Amplitude of the grinding wheel frequency f_3 for increasing number of workpieces. Sampling frequency $f_0 = 100\,\text{Hz}$

principle" for the driven mechanical parts allows one to use already-existing measured variables and to avoid additional force or torque sensors which are not only too expensive, but also introduce elasticity and reduce the reliability.

Signal-model-based evaluation of, e.g. motor currents is feasible if faults show up in a change of the oscillating behavior, like chattering. However, structure-borne noise has not generally allowed direct fault diagnosis because of too many overlapping noise sources in machine tools. However, it is used in grinding machines for detection of grinding commencement.

10

Fault detection of heat exchangers

Heat exchangers transfer heat between two or more media. They exist in a large variety of types in the chemical and power industries, in buildings and vehicles. Typical faults in these heat exchangers are leaks, e.g. by corrosion and the contamination by dirt and dissolved or suspended matter. The growth of deposits is called fouling and leads to a reduction of heat transfer. Therefore heat exchangers are usually designed with excess heat transfer surfaces of about 35% average, [10.16]. This excess design increases costs, space and weight. Remedies against fouling are chemical or mechanical mitigation techniques like filtration, additives, higher velocities, lower surface temperatures, polished surfaces. However, fouling cannot usually be avoided completely and therefore periodic cleaning will still be necessary. For more details see, e.g. [10.19], part Oc 1 and [10.17]. The detection of leaks in heat exchangers may be based on mass balances and methods described in Chapter 7. Fouling increases mainly the heat transfer coefficients, respectively the heat transfer resistance and to a minor extent the flow resistance of the media. The following sections describe some methods to detect changes of the heat transfer and some experimental results for steam-heated tubular heat exchangers with linear and parameter variable models.

10.1 Heat exchangers and their models

10.1.1 Heat exchanger types

Heat exchangers are typical apparatuses in the fields of power and chemical engineering, heating, cooling, refrigeration and air conditioning, and are part of all kinds of machines and engines. Their task is to transport heat between two or more media, e.g. liquids or gases. A large variety of types exist to meet the specific requirements with regard to temperatures, pressures, phase changes, corrosion, efficiency, weight, space and connections. Frequently used types are:

- tubular heat exchangers

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_10, © Springer-Verlag Berlin Heidelberg 2011

- plate heat exchangers

see Figures 10.1 and 10.2.

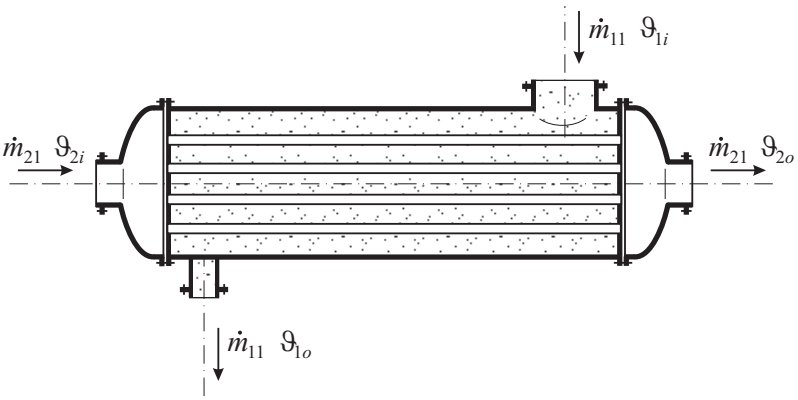


Fig. 10.1. Tubular heat exchanger [10.19]

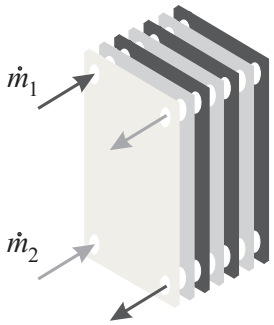


Fig. 10.2. Plate heat exchanger [10.19]

With regard to the flow direction one distinguishes counter flow, parallel flow and cross-flow. The fluids are liquids, gases or steam resulting in two media with the combinations:

- liquid–liquid
- gas–liquid
- liquid–steam (condenser, evaporator)
- gas–steam.

A type of heat exchanger widely used in the chemical process industries is that of the *shell-and-tube arrangement*, Figure 10.1. One fluid flows inside the tubes, while the other fluid flows through the shell over the outside of the tubes. To force the

outer fluid to flow across the tubes, baffles may be placed in the shell to enforce the heat transfer. The ends of the heat exchangers are called heads. According to their arrangement one or more tube passes are utilized.

Another type is *plate heat exchangers*, Figure 10.2. They consist of several metal sheets with special profiles, holes and seals. They are pressed to each other by press screws and allow a large heat transfer flow.

Cross-flow heat exchangers as in Figure 10.3, are commonly used in gas heating or cooling, like in air-conditioning units or for engine cooling of automobiles. The flow outside across the tubes is partially mixed or, if fins separate the tube bundles into passes, is unmixed.

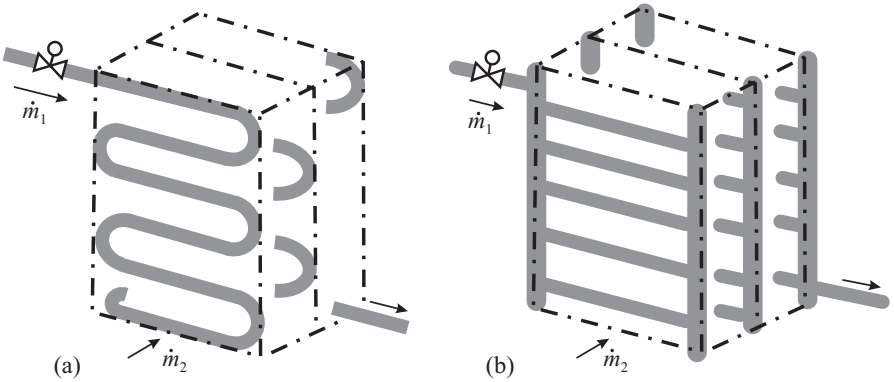


Fig. 10.3. Air–liquid cross-flow heat exchanger: a) with winding tubes; b) with parallel tubes

Within this chapter the following special symbols are used:

A	area
c_p	specific heat capacity at constant pressure
d	diameter
l	length of tubes
k	overall heat transfer coefficient
$\dot{m}$	mass flow rate
$\dot{q}$	specific heat flow $\dot{q} = \dot{Q}/A$
$\dot{Q}$	heat flow
r	vaporization heat, residual
s	tube wall thickness, Laplace variable $s = \sigma + i\omega$
v	velocity
z	tube length coordinate
α	heat transfer coefficient
ϑ	temperature (instead of T: time constant)
ρ	density

λ thermal conductivity

Subscripts:

1 primary side of heat exchanger
2 secondary side of heat exchanger
w wall
s steam
i inlet
o outlet

A bar like $\overline{\dot{m}}$ means steady state.

10.1.2 Heat exchanger models for stationary behavior

Heat-exchanging fluids are usually separated by a wall. If the heat transfer is through convection the stationary specific heat flow is, compare Figure 10.4,

$$\dot{q}_{1w} = \frac{\dot{Q}_{1w}}{A_{1w}} = \alpha_1 \left(\vartheta_{F1} - \vartheta_{w1}\right) \tag{10.1.1}$$

with A_{1w} the surface area and α_1 the heat transfer coefficient. Similarly it holds for the other side of the wall that

$$\dot{q}_{2w} = \frac{\dot{Q}_{2w}}{A_{2w}} = \alpha_2 \left(\vartheta_{w2} - \vartheta_{F2}\right) \tag{10.1.2}$$

The conductive heat transfer through the wall is (Fourier's law)

$$\dot{q}_w = \frac{\dot{Q}_w}{A_w} = \frac{\lambda}{s} \left(\vartheta_{w1} - \vartheta_{w2}\right) \tag{10.1.3}$$

with λ the thermal conductivity coefficient and s the wall thickness. In the stationary case all heat flows are identical:

$$\dot{q}_{1w} = \dot{q}_{2w} = \dot{q}_w$$

The overall heat transfer then follows if $A_{1w} = A_{2w} = A_w$ as

$$\dot{q}_{12} = \frac{\dot{Q}_{12}}{A_w} = k \left(\vartheta_{F1} - \vartheta_{F2}\right) \tag{10.1.4}$$

with the overall heat transfer coefficient

$$k = \frac{1}{\frac{1}{\alpha_{1w}} + \frac{s}{\lambda} + \frac{1}{\alpha_{2w}}} \tag{10.1.5}$$

Hence, the smallest value of α_{1w} or α_{2w} dominates the overall heat transfer.

The heat transfer through a tubular heat exchanger with mass flows $\dot{m}_1$ and $\dot{m}_2$ and specific heat capacity c_{p1} and c_{p2} is defined as

$$\dot{Q}_{12} = k\ A_w\ \Delta\vartheta_m \tag{10.1.6}$$

where ϑ_m is the average temperature difference

$$\Delta\vartheta_m = \frac{\Delta\vartheta_{la} - \Delta\vartheta_{sm}}{\ln\left(\Delta\vartheta_{la}/\Delta\vartheta_{sm}\right)} \approx \frac{1}{2}\left(\Delta\vartheta_{la} + \Delta\vartheta_{sm}\right) \tag{10.1.7}$$

and $\Delta\vartheta_{la}$ and $\Delta\vartheta_{sm}$ the larger and smaller temperature difference according to Figure 10.5.

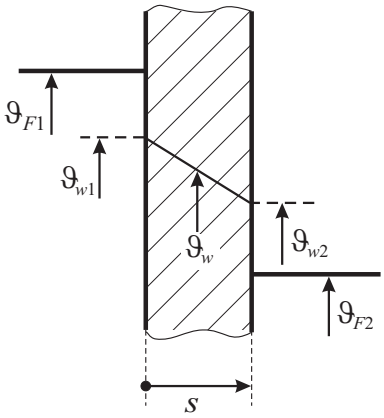


Fig. 10.4. Heat transfer through a wall

The temperature profiles in the longitudinal direction show an exponential course. It holds for example under ideal conditions for counter flow:

$$\vartheta_{22} = \vartheta_{11} - \frac{1 - \frac{\dot{m}_1 c_{p1}}{\dot{m}_2 c_{p2}}}{1 - \frac{\dot{m}_1 c_{p1}}{\dot{m}_2 c_{p2}} \exp\left[\left(\frac{1}{\dot{m}_2 c_{p2}} - \frac{1}{\dot{m}_1 c_{p1}}\right) kA\right]} \left(\vartheta_{11} - \vartheta_{21}\right) \tag{10.1.8}$$

and for parallel flow:

$$\vartheta_{22} = \vartheta_{11} - \frac{1 - \exp\left[\left(\frac{1}{\dot{m}_2 c_{p2}} - \frac{1}{\dot{m}_1 c_{p1}}\right) kA\right]}{1 - \frac{\dot{m}_1 c_{p1}}{\dot{m}_2 c_{p2}} \exp\left[\left(\frac{1}{\dot{m}_2 c_{p2}} - \frac{1}{\dot{m}_1 c_{p1}}\right) kA\right]} \left(\vartheta_{11} - \vartheta_{21}\right) \tag{10.1.9}$$

For details see, e.g. [10.7], [10.8], [10.18].

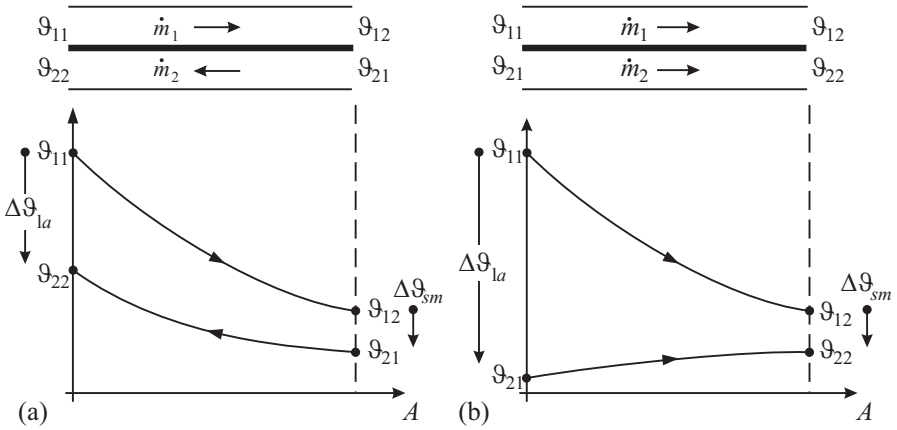


Fig. 10.5. Temperature profiles in tubular (double-pipe) heat exchangers: a) counter flow; b) parallel flow

10.1.3 Dynamic models of heated tubes

Heated tubes are basic components of many heat exchanger types. Therefore modeling the dynamic behavior of a heated tube serves as a basic element to derive and understand the temperature dynamics of heat exchangers. In many cases it is of interest how the output temperature $\vartheta_{1o}(t)$ depends dynamically on the input temperature $\vartheta_{1i}(t)$, the velocity $v_1(t)$ and the specific heat flow $\dot{q}_{2w}(t)$, see Figure 10.6. The description below follows [10.1], [10.9], [10.11], [10.18]:

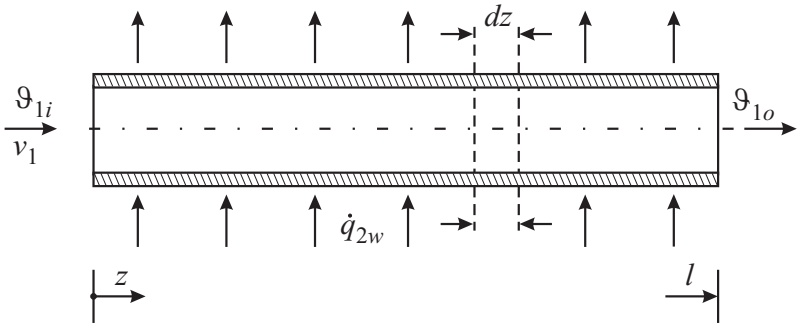


Fig. 10.6. Scheme of a heated tube

a) Heated tube with distributed parameters

It is assumed that the geometric dimensions and the specific heat flow $\dot{q}_{2w}(z)$ are constant along the length coordinate z and that a turbulent flow holds internally, and that there is an ideal mixture perpendicular to the z-axis.

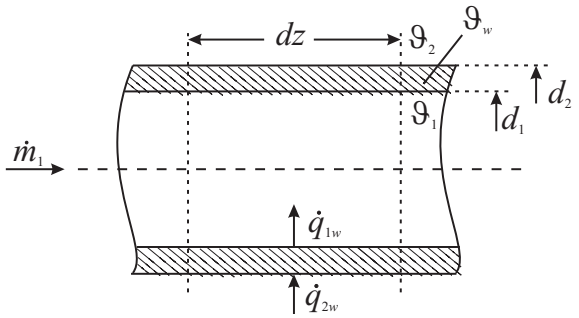


Fig. 10.7. Tube element

As the tube is a distributed parameter system where the fluid temperature $\vartheta_1(z,t)$ depends on the location z and the time t an infinitesimally small element is considered, see Figure 10.7. It consists of two heat stores, the inner fluid and the tube wall. Therefore two balance equations are stated:

- enthalpy balance equation for the flowing fluid:

$$\frac{Dh_1(z,t)}{Dt} = \frac{1}{A_1\rho_1}\frac{\partial \dot{Q}_{1w}}{\partial z}\ ;\ d\dot{Q}_{1w} = \dot{q}_{1w}\pi d_1 dz \tag{10.1.10}$$

$$\frac{\partial h_1(z,t)}{\partial t} + v_1(t)\frac{\partial h_1(z,t)}{\partial z} = \frac{\pi d_1}{A_1\rho_1}\dot{q}_{1w}(z,t) \tag{10.1.11}$$

- heat balance of the tube wall element:

$$\dot{q}_{2w}\ \pi\ d_2\ dz - \dot{q}_{1w}\ \pi\ d_1\ dz = A_w\ \rho_w\ c_w\ \frac{\partial \vartheta_w}{\partial t} \tag{10.1.12}$$

where $A_w = \pi\left(d_2^2 - d_1^2\right)$ is the cross-section area of the tube wall.

After assuming small changes of the temperature $\Delta\vartheta(t)$ and other variables and introducing the heat transfer equations

$$\dot{q}_{1w} = \alpha_{1w}(\vartheta_w - \vartheta_1) \tag{10.1.13}$$

$$\dot{q}_{2w} = \alpha_{2w}(\vartheta_2 - \vartheta_w) \tag{10.1.14}$$

with Nusselt's law for the heat transfer coefficient

$$\begin{aligned} \alpha_{1w} &= \alpha_{1w}\left(\frac{v_1}{\overline{v}_1}\right)^m \\ \Delta\alpha_{1w} &= \overline{\alpha}_{1w}\ m\ \frac{\Delta v_1}{\overline{v}_1} \end{aligned} \tag{10.1.15}$$

($m = 0.8$ for turbulent flow, $\overline{v}$ means steady-state value) and for the enthalpy

$$\Delta h_1 = c_{p1}\ \Delta\vartheta_1 \tag{10.1.16}$$

it follows:

$$\begin{aligned}\frac{\partial \vartheta_1(z,t)}{\partial t} + \overline{v}_1 \frac{\partial \vartheta_1(z,t)}{\partial z} &= \frac{1}{T_F}\left(\Delta\vartheta_w(z,t) - \Delta\vartheta_1(z,t)\right) \\ &\quad - \frac{1}{T_F}\left(\overline{\vartheta}_w - \overline{\vartheta}_1\right)(m-1)\frac{\Delta v_1}{\overline{v}_1}\end{aligned} \tag{10.1.17}$$

$$\begin{aligned}\frac{\partial \vartheta_w(z,t)}{\partial t} &= \frac{1}{\alpha_{2w}\ T_{w2}}\Delta\dot{q}_{2w}(t) - \frac{1}{T_{w2}}\Delta\vartheta_w(z,t) \\ &\quad + \frac{1}{T_{w1}}\left(\Delta\vartheta_1(z,t) - \Delta\vartheta_w(z,t)\right) \\ &\quad - \frac{1}{T_{w1}} m \left(\overline{\vartheta}_w - \overline{\vartheta}_1\right)\frac{\Delta v_1(t)}{\overline{v}_1}\end{aligned} \tag{10.1.18}$$

Herewith three parameters are defined:

$$T_F = \frac{d_1\ \rho_1\ c_{p1}}{4\ \overline{\alpha}_{1w}} \qquad \text{(fluid time constant)} \tag{10.1.19}$$

$$T_{w1} = \frac{A_w\ \rho_w\ c_w}{\pi\ d_1\ \overline{\alpha}_{1w}} \qquad \text{(inner tube wall time constant)} \tag{10.1.20}$$

$$T_{w2} = \frac{A_w\ \rho_w\ c_w}{\pi\ d_2\ \overline{\alpha}_{2w}} \qquad \text{(external tube wall time constant)} \tag{10.1.21}$$

(10.1.17) and (10.1.18) are now first Laplace transformed with regard to the time t into the s-domain and then with regard to the location z in the ζ-domain. Laplace backtransformation from the ζ-domain into the z-domain with setting of $\Delta\vartheta_1(z = 0, s) = \Delta\vartheta_{1i}(s)$; $\Delta\vartheta_1 i(z = l, s) = \Delta\vartheta_{1o}(s)$ and elimination of the tube wall temperature leads to the following three transfer functions, [10.9]:

$$G_\vartheta(s) = \frac{\Delta\vartheta_{1o}(s)}{\Delta\vartheta_{1i}(s)} = e^{-T_t s}\ e^{-\kappa_F \frac{T_{w1}s+\eta}{T_{w1}s+\eta+1}} \tag{10.1.22}$$

$$\begin{aligned}G_q(s) &= \frac{\Delta\vartheta_{1o}(s)}{\Delta\dot{q}_{2w}(s)} \\ &= \frac{\left(\overline{\vartheta}_w - \overline{\vartheta}_1\right)d_2}{\dot{\overline{q}}_{2w} d_1} \frac{1/\eta}{T_{w2}\ T_F\ s^2 + \left(T_F\ \frac{1+\eta}{\eta} + T_{w2}\right)s + 1}\ (1 - G_\vartheta(s))\end{aligned} \tag{10.1.23}$$

$$\begin{aligned}G_v &= \frac{\Delta\vartheta_{1o}(s)}{\Delta v_1(s)} \\ &= -\frac{\left(\overline{\vartheta}_w - \overline{\vartheta}_1\right)}{\overline{v}_1} \frac{\frac{1}{\eta} + (1-m)(1+T_{w1}s)}{T_{w2}\ T_F\ s^2 + \left(T_F \frac{1+\eta}{\eta} + T_{w2}\right)s + 1}\ (1 - G_\vartheta(s))\end{aligned} \tag{10.1.24}$$

Three new parameters are:

$$\kappa_F = \frac{4\,\alpha_{1w}\,l}{d_1\,\rho_1\,c_{p1}\,v_1} \qquad \text{(fluid parameter)} \qquad (10.1.25)$$

$$\eta = \frac{\alpha_{2w}\,d_2}{\alpha_{1w}\,d_1} \qquad \text{(convective heating parameter)} \qquad (10.1.26)$$

$$T_t = l/v \qquad \text{(dead time)} \qquad (10.1.27)$$

and additionally from (10.1.20):

$$T_{w1} = \frac{\left(d_2^2 - d_1^2\right)\,\rho_w\,c_w}{4\,\alpha_{1w}\,d_1} \qquad \text{(inner tubewall time constant)} \qquad (10.1.28)$$

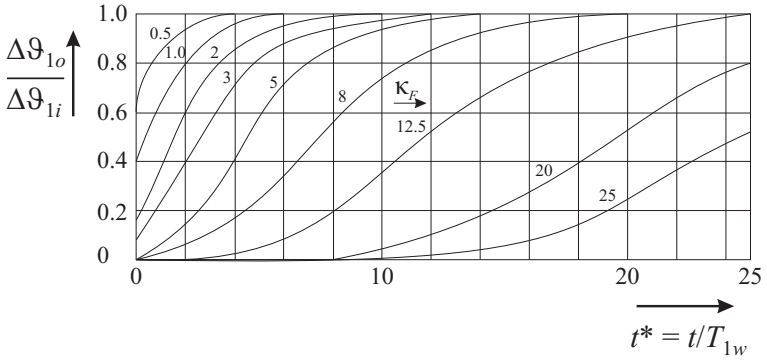


Fig. 10.8. Transient functions for $G_\vartheta(s)$ with κ_F as parameter (heating through radiation, $\eta = 0$) and assuming $T_t = 0$, [10.18]

The parameters κ_1 and η define the form of the resulting frequency response loci or the transient functions and T_{w1} and T_{w2} the frequency or time scale. Figures 10.8 and 10.9 show the corresponding transient functions and frequency response loci. The temperature behavior is for small κ_1 (e.g. short tube length), approximately a step plus a first-order lag, and tends to a higher-order lag behavior for large κ_1 (large tube length). Both the heating/temperature and the velocity/temperature behavior show approximately a second-order lag behavior, however, with opposite sign, see also Figure 10.10. Table 10.1 shows typical ranges of the characteristic parameters for heat exchangers.

b) Simplified models of heated tubes

Approximation of the transcendental transfer functions with rational transfer function and time delay allows simpler expressions, as was shown in [10.1], [10.9], [10.15]:

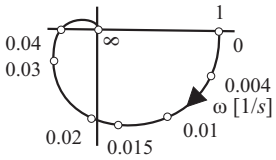
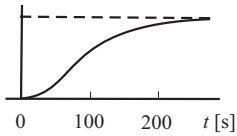
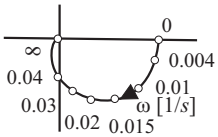
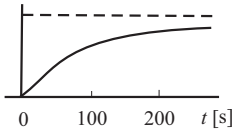
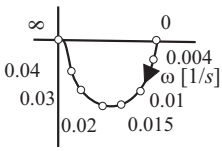
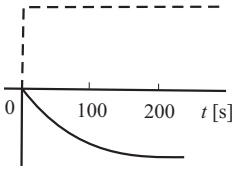
	frequency response	transient function
$\frac{\Delta\vartheta_{1o}(s)}{\Delta\vartheta_{1i}(s)}$		
$\frac{\Delta\vartheta_{1o}(s)}{\Delta\dot{q}_{1w}(s)}$		
$\frac{\Delta\vartheta_{1o}(s)}{-\Delta v_1(s)}$		

Fig. 10.9. Frequency responses and transient responses of a steam superheater, [10.18]

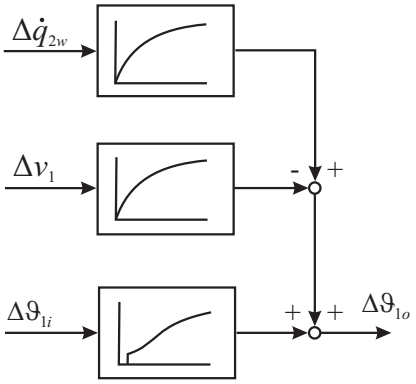


Fig. 10.10. Signal flow diagram of a heated tube for small changes of variables

Table 10.1. Characteristic parameters of heated tubes, [10.9]

primary flow 1 (inner tube)	liquid $0.1 < \kappa_F < 20$ $0.2 < T_{w1}/T_F < 0.7$	gases, steam $0.1 < \kappa_F < 20$ $20 < T_{w1}/T_F < 100$
secondary flow 2 (outer tube)	gas $0 < \eta < 0.2$ (water/air heater)	gas, steam $0 < \eta < 1$ (steam superheater)
	liquid, condensing steam $0.5 < \eta < 3$ (steam/water heat exchanger)	liquid $0.5 < \eta < 10$

$$\widetilde{G}_\vartheta(s) = \frac{\Delta\vartheta_{1o}(s)}{\Delta\vartheta_{1i}(s)} = \left(a + \frac{b}{1 + T_b\, s}\right)^n \; e^{-T_t s} \tag{10.1.29}$$

$$a = e^{-\Delta\kappa_F}$$

$$b = e^{-\Delta\kappa_F \frac{\eta}{1+\eta}} - a$$

$$T_b = \frac{\Delta\kappa_F}{(1+\eta)^2}\,\frac{a+b}{b}\,T_{w1}$$

$$n = \frac{\kappa_1}{\Delta\kappa_F} \text{ with } \Delta\kappa_1 \le 1.5, n = 1, 2, 3, \ldots$$

This means that (10.1.22) was approximated for small $\Delta\kappa_F$ by rational transfer functions. For larger κ_F they are connected in series. The other two transfer functions result in

$$\widetilde{G}_q(s) = \frac{\Delta\vartheta_{1o}(s)}{\Delta\dot{q}_{2w}(s)} = \frac{G_q(0)}{(1 + T_1 s)\,(1 + T_2 s)} \tag{10.1.30}$$

$$T_1 = \frac{1}{\eta}\,(1 - \psi)\,T_{w1}$$

$$T_2 = \frac{1}{\eta}\,T_{w1}$$

$$\widetilde{G}_v(s) = \frac{\Delta\vartheta_{1o}(s)}{\Delta v_1(s)} = \frac{G_v(0)}{(1 + T_1 s)}\left\{\frac{\frac{1}{\eta}}{1 + T_2 s} + (1 - m)\right\} \tag{10.1.31}$$

The gains are

$$G_\vartheta(0) = \psi = e^{-\kappa_1 \frac{\eta}{1+\eta}} \qquad (\text{if } c_{p1i} = c_{p1o})$$

$$G_q(0) = \frac{\overline{\vartheta}_w - \overline{\vartheta}_1}{\overline{\dot{q}}_{2w}}\,\frac{d_2}{d_1}\,\frac{1}{\eta}(1 - \psi)$$

$$G_v(0) = -\frac{\overline{\vartheta}_w - \overline{\vartheta}_1}{\overline{v}_1}\left(\frac{1}{\eta} + 1 - m\right)(1 - \psi)$$

These approximations hold for liquid flows in the inner tube with ($0.2 < T_{w1}/T_F < 0.7$) and liquids and condensing steam for the secondary flow ($0.5 < \eta < 3$). For other parameter combinations see [10.9]. For simplified dynamic models of cross flow heat exchangers see [10.14].

The output temperature is usually measured by a *temperature sensor*, e.g. a resistance thermometer or a thermocouple. Therefore the transfer functions for the outlet temperature of the heated tube have to be multiplied by the transfer function:

$$G_{ts} = \frac{\Delta\vartheta_{1s}(s)}{\Delta\vartheta_{1o}(s)} = \frac{K_{ts}}{1 + T_{ts}s} \tag{10.1.32}$$

$$T_{ts} = \frac{d_{ts}\rho_{ts}c_{ts}}{4\alpha_{1ts}} \tag{10.1.33}$$

with d_{ts} diameter, ρ_{ts} density, c_{ts} specific heat coefficient, α_{1ts} heat transfer coefficient of the sensor. The sensor dynamics may have a significant influence on the overall dynamic behavior.

10.2 Fault detection for static behavior

10.2.1 Static models of heat exchangers

A general heat exchanger according to Figure 10.11 is considered with primary fluid 1 and secondary fluid 2, both of which are in a stationary operation state. It is assumed that the following variables are measurable:

$\dot{m}_1, \dot{m}_2$ mass flows
$\vartheta_{1i}, \vartheta_{2i}$ input temperatures
$\vartheta_{1o}, \vartheta_{2o}$ output temperatures

A heat balance then leads to

$$\overline{\dot{Q}}_1 = \overline{\dot{m}}_1\, c_{p1}\, \left(\overline{\vartheta}_{1i} - \overline{\vartheta_{1o}}\right) = \overline{\dot{m}}_2\, c_{p2}\, \left(\overline{\vartheta}_{2o} - \overline{\vartheta}_{2i}\right) - \dot{Q}_l \tag{10.2.1}$$

where $\dot{Q}_l$ is the heat flow loss to the environment and c_p is the specific heat at constant pressure. The overall heat transfer coefficient of a heat exchanger is, see (10.1.6)

$$\overline{k}_{HE} = \frac{\overline{\dot{Q}}_1}{A\ \Delta\overline{\vartheta}_m} \left[\frac{W}{m^2\ K}\right] \tag{10.2.2}$$

where A is the exchange surface area and $\Delta\vartheta_m$ the mean temperature difference. $\Delta\vartheta_m$ depends on the direction of liquid flows, e.g. either parallel, reverse or cross-flow, and is e.g. for reverse flow, compare (10.1.7) and Figure 10.5,

$$\Delta\overline{\vartheta}_{la} = \overline{\vartheta}_{1i} - \overline{\vartheta}_{2o} \quad ; \quad \Delta\overline{\vartheta}_{sm} = \overline{\vartheta}_{1o} - \overline{\vartheta}_{2i}$$

$$\Delta\overline{\vartheta}_m = \frac{\Delta\overline{\vartheta}_{la} - \Delta\overline{\vartheta}_{sm}}{ln\left(\Delta\overline{\vartheta}_{la}/\Delta\overline{\vartheta}_{sm}\right)} \tag{10.2.3}$$

The overall heat transfer coefficient of a thin tube wall is according to (10.1.5)

$$\overline{k} = \frac{1}{\frac{1}{\alpha_1} + \frac{s}{\lambda} + \frac{1}{\alpha_2}} \left[\frac{W}{m^2\ K}\right] \tag{10.2.4}$$

where α are the heat transfer coefficients, λ the thermal conductivity coefficient and s the wall thickness.

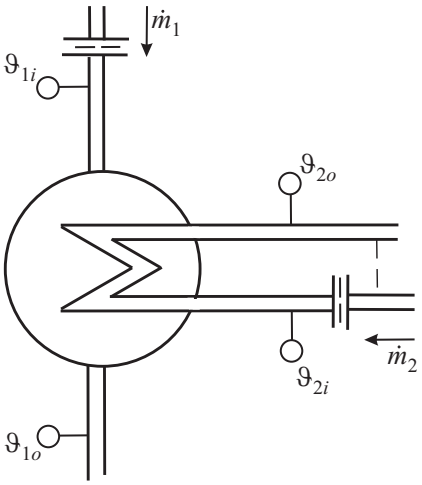


Fig. 10.11. General scheme of a heat exchanger

10.2.2 Fault-detection methods

a) Parity equation

The simplest method for an overall fault detection is to calculate a residual of the heat balance (10.2.1)

$$r(k) = \overline{\dot{m}}_1(k)\ \left(\overline{\vartheta}_{1i}(k) - \overline{\vartheta}_{1o}(k)\right) - \overline{\dot{m}}_2(k)\ \left(\overline{\vartheta}_{2o}(k) - \overline{\vartheta}_{2i}(k)\right) \tag{10.2.5}$$

where measurements are made at different discrete times $k = t/T_0$ under steady-state conditions. The residual corresponds then to the heat loss

$$r(k) = \overline{\dot{Q}}_l(k)/c_p \tag{10.2.6}$$

This residual will change in the case of faults like defective insulation, any faults of the six sensors, leaks in flow 1 or flow 2.

However, these faults are not isolable. Also contamination on one or two sides is not directly detected, as long as the heat balance is satisfied ($\dot{Q}_l$ may change a bit). A further disadvantage is the requirement for six sensors. In special cases one sensor can be saved, e.g. if $\overline{\vartheta}_{1o} = \overline{\vartheta}_{2i}$ can be assumed.

b) Characteristic quantity

Another overall feature is the mean heat transfer coefficient (10.2.2)

$$k_{HE}(k) = \frac{\overline{\dot{Q}}(k)}{A\ \Delta\overline{\vartheta}_m(k)} = \frac{1}{A\ \Delta\overline{\vartheta}_m(k)}\ \overline{\dot{m}}_1(k)\ c_{p1}\ \left(\overline{\vartheta}_{1i}(k) - \overline{\vartheta}_{1o}(k)\right) \tag{10.2.7}$$

Comparing this quantity with a nominal value

$$\Delta k_{HE}(k) = k_{HE}(k) - k_{HEnom} \tag{10.2.8}$$

indicates faults in the insulation, leaks at both sides, contamination on both sides, faults of three sensors.

However, also these faults are not isolable. The number of sensors is still high: one mass flow and four temperature sensors.

c) Parameter estimation

As contamination (fouling) within and outside of the tubes increases the heat transfer coefficients it would be advantageous to estimate these parameters directly. However, they are hidden in the overall transfer coefficient (10.2.4).

The heat transfer coefficients depend on the mass flow rate according to Nusselt's law:

$$\alpha_1 = c_{\alpha 1}\ \overline{\dot{m}}^{\beta_1} \qquad \alpha_2 = c_{\alpha 2}\ \overline{\dot{m}}^{\beta_2} \tag{10.2.9}$$

If for example in a shell-tube heat exchanger or cross-flow heat exchanger flow 1 streams through tube 1 with turbulent flow it holds that $\beta_1 \approx 0.8$, and if flow 2 streams perpendicularly through a bundle of tubes it is $\beta_2 \approx 0.6$. Then α_1 and α_2 depend differently on the mass flows. Using (10.2.4) and writing it in the form

$$\frac{1}{\alpha_1} + \frac{s}{\lambda} + \frac{1}{\alpha_2} = \frac{1}{k} \tag{10.2.10}$$

leads to

$$\begin{aligned} \frac{1}{c_{\alpha 1}}\ \overline{\dot{m}}_1^{\beta_1} + \frac{1}{c_{\alpha 2}}\ \overline{\dot{m}}_2^{\beta_2} &= \frac{1}{\overline{k}} - \frac{s}{\lambda} = y \\ \alpha_1\ \overline{\dot{m}}_1^{\beta_1} + \alpha_2\ \overline{\dot{m}}_2^{\beta_2} &= \frac{1}{\overline{k}} - \frac{s}{\lambda} = y \end{aligned} \tag{10.2.11}$$

Now it is assumed that $\overline{k} = \overline{k}_{HE}$ is known by applying (10.2.2) where the heat flow $\dot{Q}_1$ follows from (10.2.1) by measurement of $\dot{m}_1$, ϑ_{1i} and ϑ_{1o} and $\Delta\vartheta_m$ follows from (10.2.3) by additional measurements of ϑ_{2o} and ϑ_{2i}. Also s and λ are known from construction data. (10.2.11) can be written as

$$y = \boldsymbol{\psi}^T \, \boldsymbol{\Theta} \tag{10.2.12}$$

with

$$\boldsymbol{\psi}^T = \left[\overline{\dot{m}}_1^{\beta_1} \; \overline{\dot{m}}_2^{\beta_2}\right]$$
$$\boldsymbol{\Theta}^T = [a_1 \; a_2]$$

The parameters a_1 and a_2 can be estimated by using the least-squares parameter estimation method. Introducing the equation error $e(k)$ in (10.2.11) yields

$$y(k) = \boldsymbol{\psi}^T(k) \, \boldsymbol{\Theta} + e(k) \tag{10.2.13}$$

The minimization of the sum of squared errors leads to the least-squares estimate

$$\boldsymbol{\Theta}(N) = \left[\boldsymbol{\Psi}^T \, \boldsymbol{\Psi}\right]^{-1} \, \boldsymbol{\Psi}^T \, \boldsymbol{y} \tag{10.2.14}$$

where $\boldsymbol{\Psi}$ contains different mass flow measurements $\overline{\dot{m}}_1(k)$ and $\overline{\dot{m}}_2(k)$ in a steady state. However, a good condition of the equation system can only be expected for wide ranges of $\dot{m}_1$ and $\dot{m}_2$ measurements.

The estimated parameters $\hat{a}_1$ and $\hat{a}_2$ in $\hat{\boldsymbol{\Theta}}$ allow the calculation of $\hat{c}_{\alpha 1}$ and $\hat{c}_{\alpha 2}$, see (10.2.9). Decreasing values

$$\Delta\hat{c}_{\alpha i} = \hat{c}_{\alpha i} - c_{\alpha nom} \quad i = 1.2 \tag{10.2.15}$$

then enable one to detect contamination of the heat transfer surfaces under the condition that sensor faults can be excluded, because they influence k_{HE}. The parameter estimation is also possible if only one of the mass flows is changed. However, this method requires that four temperatures and two mass flows have to be measured.

d) Required instrumentation

The different possibilities for fault detection of heat exchangers with static models have shown that two mass flow sensors and four temperature sensors are required. If the considered heat exchanger is an important process part in the plant, then this instrumentation is usually available. However, if this sensor equipment is not completely implemented at the heat exchanger, one has to use other sensors within the plant. *Flow rates* may be measured at other positions or may be calculated based on mass flow balances. A further alternative is a model-based reconstruction of the flow rate with:

i) *pump-drive-unit model:* The flow rate is calculated based on measurements of current, speed and/or pressure difference, see Section 6.1.2
ii) *valve model:* The flow rate is calculated from measurement of valve position (resp. effective opening area) and pressure difference across the valve, see Section 5.2.

Frequently the fluid outlet temperature of one fluid, e.g. ϑ_{1o}, is controlled by a feedback controller manipulating either its inlet temperature ϑ_{2i} or the mass flow rate $\dot{m}_2$ of the other fluid or its inlet temperature ϑ_{2i}, compare Figure 10.11. Then the position U of the actuator can be used as an approximation of the corresponding manipulated variable. If U leaves a normal operating range, or if it reaches the actuation limit and a permanent offset of the control deviation arises, the heat transfer in the heat exchanger may be faulty, if other faults (sensors, actuators) can be excluded, compare [10.12], Chapter 12.

In the case of *steam-heated heat exchangers* with saturated steam one temperature sensor may be saved, because the transferred heat flow is

$$\dot{Q}_1 = \dot{m}_1 \, r(\vartheta_{1i})$$

where r is the specific vaporization heat. ϑ_{1o} has not to be measured if the condensate valve functions properly, avoiding condensate storage in the heat exchanger. The measurement of ϑ_{1i} may be replaced by a pressure measurement p_{1i}.

10.3 Fault detection for a steam/water heat exchanger with dynamic models and parameter estimation

An industrial-size steam-heated heat exchanger, see Figure 10.12, is considered which is part of a pilot plant, [10.6], [10.13]. This plant consists of an electric-powered steam generator, a steam/condensate circulation (circuit 1), a water circulation (circuit 2) and a cross-flow heat exchanger to transport the heat from water to air. As inputs and outputs of the considered heat exchanger the following variables are measured:

$\dot{m}_s$ mass flow of the steam
$\dot{m}_1$ mass flow of the liquid fluid (water)
ϑ_{1i} inlet temperature of the liquid fluid
ϑ_{1o} outlet temperature of the liquid fluid

The fluid outlet temperature ϑ_{1o} is considered as an output variable, the other three measured variables as input variables.

10.3.1 Fault detection with linear dynamic models and parameter estimation

To model the dynamic behavior, the heat exchanger is subdivided into the tubular section, the water head, a transport delay and the temperature sensor, see [10.10]. The dynamic equations for a heated tube are obtained from section 10.1.3b). In addition, balance equations are stated for the steam space and the shell tube. Then the equation system is linearized around the operating point. For the steam flow as input one obtains for example the approximate transfer function

$$\widetilde{G}_{s\vartheta}(s) = \frac{\Delta\vartheta_{1o}(s)}{\Delta\dot{m}_s(s)} = \frac{K_s}{(1 + T_{1s}s)(1 + T_{2s}s)} e^{-T_{ts}s} \tag{10.3.1}$$

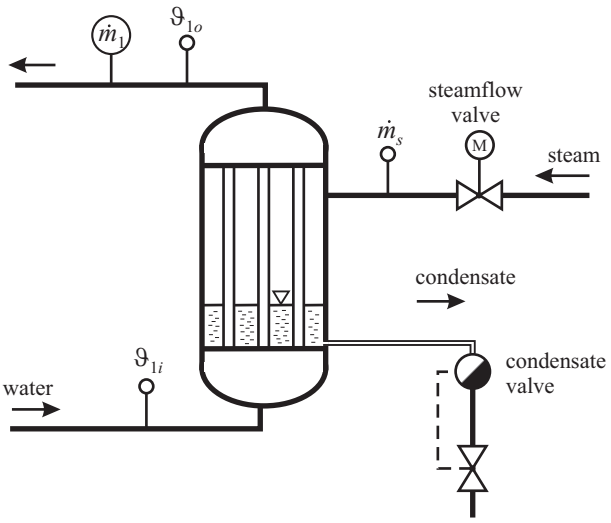


Fig. 10.12. Tubular heat exchanger and measured variables

with

$$\left.\begin{aligned} K_s &= \frac{r}{\dot{m}_1 c_1}; \quad T_{1s} = \frac{1}{v_1}\left[1 + \frac{A_w \rho_w c_w}{A_1 \rho_1 c_1}\right] \\ T_{2s} &= \frac{A_w \rho_w c_w}{\alpha_{w1} U_1} \frac{1}{\left[1 + \frac{A_w \rho_w c_w}{A_1 \rho_1 c_1}\right]} \end{aligned}\right\} \tag{10.3.2}$$

The various symbols are:

A	cross-sectional area
c	specific heat capacity
$m, \dot{m}$	mass, mass flow rate
r	evaporation heat
U	periphery of one tube
v	velocity in the tube
α	heat transfer coefficient
ϑ	temperature
ρ	density

Subscripts:

1	fluid (water)
s	steam
w	wall
i	inlet
o	outlet

In this case three parameter estimates compare to 10 process coefficients. Therefore it is not possible to determine all process coefficients uniquely. By assuming some of the process coefficients to be known, however, the following process coefficients and process coefficient combinations can be determined:

$$\left.\begin{aligned} \alpha_{w1} &= \frac{A_1\rho_1 c_1}{T_{2s}U_1}\left[1-\frac{1}{T_{1s}v_1}\right] \\ A_w\rho_w c_w &= T_{1s}\dot{m}_1 c_1 - A_1\rho_1 c_1 \\ r &= K_s\dot{m}_1 c_1 \end{aligned}\right\} \tag{10.3.3}$$

The three parameters $\hat{K}_s$, $\hat{T}_{1s}$ and $\hat{T}_{2s}$ are determined by experiments based on transient function measurements of the fluid outlet temperature ϑ_{1o} due to changes of the input variables ϑ_{1i}, $\dot{m}_S$ and $\dot{m}_1$ in the direction of decreasing temperature ϑ_{So}. The operating point was

$$\dot{m}_1 = 3000 \text{ kg/h};\ \dot{m}_S = 50 \text{ kg/h};\ \vartheta_{1i} = 60^\circ \text{ C}; \vartheta_{1o} \approx 70^\circ \text{ C}$$

As the sampling time $T_0 = 500$ ms was selected. The time period of one experiment was 360 s, so that 720 samples were taken. For the parameter estimation the method of total least squares in a recursive form was applied by using a digital state-variable filter for the determination of the derivatives. The determination of the normal state (training phase) was based on 60 transients for each transfer function. The 30 transients were carried out for each transfer function and for each of four artificially generated faults:

F1: air (inert gas) in the steam space
F2: open condensate valve
F3: closed condensate valve
F4: plugged tube

(Altogether 540 experiments were carried out, lasting about 150 operation hours.)

Figure 10.13 shows (a) one measured transient function and (b) the corresponding time history of the parameter estimates. A good convergence of the parameter estimates was obtained in all cases. A verification of the measured and the calculated transient functions shows a very good agreement. In Table 10.2 the parameter estimates are given for $\widetilde{G}_{s\vartheta}$. Table 10.3 indicates that for each of the four faults different changes in the parameter estimates are obtained. Corresponding results were obtained for changes in the fluid inlet temperature $\widetilde{G}_{\vartheta\vartheta}$, and changes in fluid flow $\widetilde{G}_{1\vartheta}$.

Based on the parameter estimates $\hat{K}_s$, $\hat{T}_{1s}$, $\hat{T}_{2s}$ the obtainable process coefficients due to (10.3.3) can also be calculated. However, in this case the changes in the calculated process coefficients can only partially be explained by the physical effects of the faults. One reason is that the investigated faults are not directly mapped by the obtainable process coefficients. In addition, the calculated process coefficients are obviously rather sensitive to changes in the process parameter estimates and to the

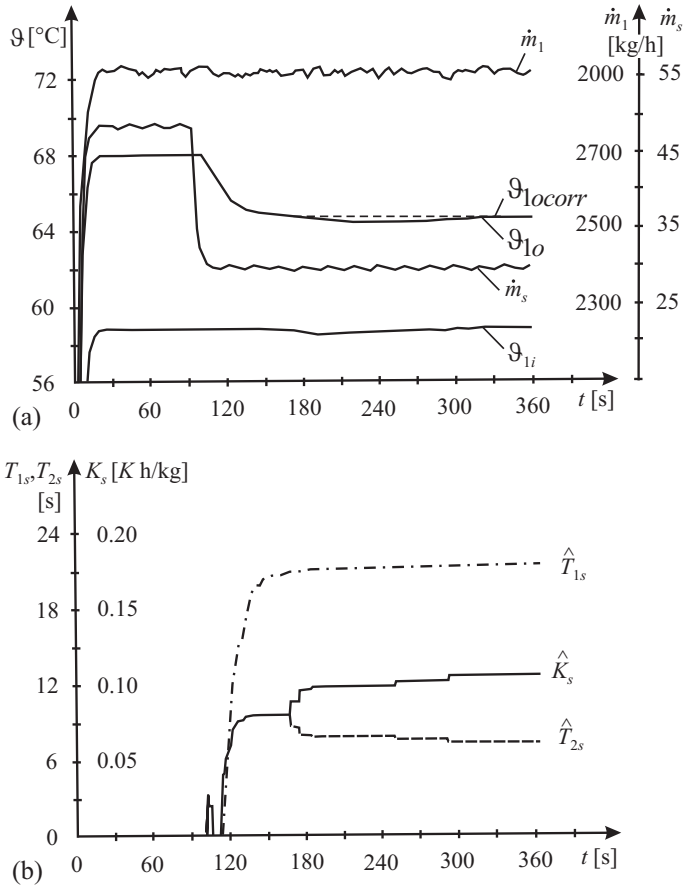


Fig. 10.13. Results for a change in the steam flow $\Delta\dot{m}_s$: a) Measured transient functions for a steam flow change; b) Parameter estimates from transient function

Table 10.2. Parameter estimates for steam flow changes

fault	mean standard deviation	$\hat{K}_s$ [Kh/kg]	$\hat{T}_{1s}$ [s]	$\hat{T}_{2s}$ [s]
none	μ	0.1708	12.38	7.21
	σ	0.0032	1.63	1.07
F1	μ	0.1896	7.26	7.26
	σ	0.0072	0.73	0.73
F2	μ	0.1268	7.26	7.26
	σ	0.0037	0.35	0.35
F3	μ	0.1899	13.89	3.81
	σ	0.0042	0.82	0.44
F4	μ	0.1689	13.65	6.01
	σ	0.0032	1.50	0.81

Table 10.3. Changes of the parameter estimates (symptoms) of Table 10.2 for $\tilde{G}_{S\vartheta}(s)$. $+$, $++$ small, large increase, $-$, $--$ small, large decrease, 0 no change

fault	$\hat{K}_s$	$\hat{T}_{1s}$	$\hat{T}_{2s}$
F1	$-$	$--$	0
F2	$--$	$--$	$+$
F3	$+$	$+$	$--$
F4	0	$+$	$-$

values of the coefficients which have to be assumed as known. This is a known fact for processes with heat transfer. Therefore it is recommended that the fault detection for this type of process, which is a higher-order distributed parameter system and is approximated by lower-order lumped models, be based on the model parameter estimates such as gain and time constants. This means that a detailed theoretical modeling is not necessary. The case study has shown that the considered four faults could be detected by using patterns of changes according to Table 10.3. They are strongly isolable. This holds for all three transfer functions, [10.6]. Hence it is sufficient to use only one transfer function. The most significant differences were obtained for steam flow changes, $\widetilde{G}_{s\vartheta}$. The use of static models only allows the recognition of changes from the normal state. Hence, a detailed fault detection and diagnosis is in this case only possible by applying dynamic models, compare [10.17]. The required measurements are two mass flows, two temperatures and one (steam) pressure.

10.3.2 Fault detection with parameter variable local linear dynamic models

As the behavior of heat exchangers depends strongly on the flow-rates the static and dynamic behavior is nonlinear for changing flow-rates. In order to develop fault-detection methods which are applicable over a large operating range local linear neuronal net models of the type LOLIMOT were used to describe first the nominal behavior. This was applied to the steam/water heat exchanger used also for Section 10.3.1, [10.2], [10.5]. By using the LOLIMOT identification method dynamic models of the water outlet temperature ϑ_{1o} in dependence on the water volume flow $\dot{V}_1$, steam mass flow $\dot{m}_s$ and inlet temperature ϑ_{1i} were determined by simultaneous wide-range excitation of the two flows with amplitude-modulated PRBS, [10.4]. This resulted in 10 local linear models in dependence on the water flow. Using a sample time of $T_0 = 1$ s a second-order dynamic model was sufficient:

$$\begin{aligned}\vartheta_{1o}(k) = &-a_1(z)\vartheta_{1o}(k-1) - a_2(z)\vartheta_{1o}(k-2)\\ &+ b_{11}(z)\dot{m}_s(k-1) + b_{12}(z)\dot{m}_s(k-2)\\ &+ b_{21}(z)\dot{V}_1(k-1) + b_{31}(z)\vartheta_{1i}(k-1) + c_0(z)\end{aligned} \tag{10.3.4}$$

where the parameters depend on the operating point $z = \dot{V}_1$ (volume flow rate)

$$a_\nu(\dot{V}_1) = \sum_{j=1}^{10} a_\nu \Phi_j(\dot{V}); \quad b_{\nu\mu}(\dot{V}_1) = \sum_{j=1}^{10} b_{\nu\mu}(\dot{V})$$
$$c_0(\dot{V}_1) = \sum_{j=1}^{10} c_0 \Phi_j(\dot{V}) \tag{10.3.5}$$

where Φ_j is the weighting function within LOLIMOT.

Figure 10.14 shows the resulting stationary outlet temperature in dependence on the two flows. The identified models then allow one to extract three gains and one dominant time constant, partially depicted in Figure 10.15. The operating-point dependence is especially strong for low water flow rates. Static gains and the time constant change with about a factor of four.

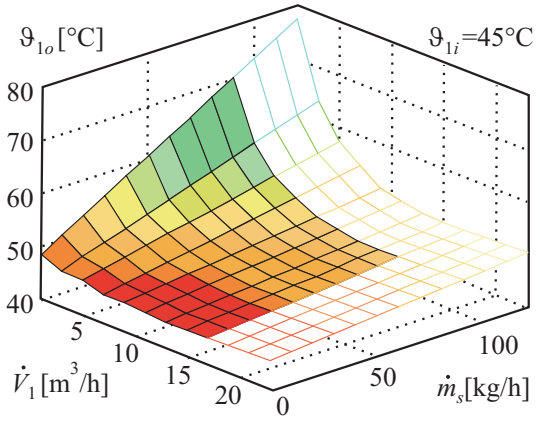


Fig. 10.14. Heat exchanger water outlet temperature static map in dependence on water and steam flow

Based on the local linear models several features can be extracted for fault detection:

$K_{\dot{v}_1}$	static gain for $\Delta\dot{V}_1$
$K_{\dot{m}_s}$	static gain for $\Delta\dot{m}_s$
K_{ϑ_1}	static gain for $\Delta\vartheta_{1i}$
T_{1o}	time constant
c_0	static offset
$r_{\vartheta_1}(k) = \vartheta_{1o}(k) - \vartheta_{1onom}(k)$	output residual

the following faults were then inserted:

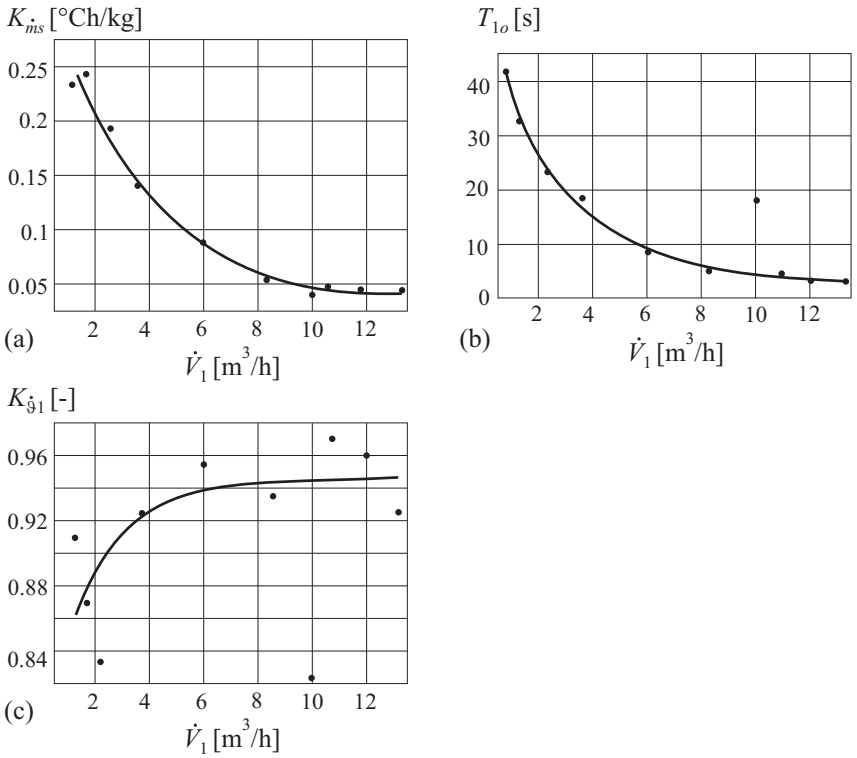


Fig. 10.15. Static gains and time constant for the water outlet temperature in dependence on the water flow rate

F1: leak in the heat exchanger (by opening a bypass)
F2: condensate valve stuck open
F3: condensate valve stuck closed
F4: leak in the vacuum pump (too little pump delivery)
F5: inert gas (air) in the steam space
F6: gain fault in sensor ϑ_{1i}
F7: gain fault in sensor $\dot{V}_1$
F8: gain fault in sensor ϑ_{1o}
F9: gain fault in sensor $\dot{m}_s$
(gain faults: factor 1.2)

For the experiments without and with inserted faults the process was excited by changing the inputs $\dot{V}_1$, $\dot{m}_s$ and ϑ_{1i} with a PRBS around the operating points with water volume flows $\dot{V}_1 = 4, 8$ and $13\,\mathrm{m}^3/\mathrm{h}$. The parameters were then estimated with recursive least-squares estimation (RLS) and forgetting memory by applying the same second-order models as (10.3.4), [10.3]. Table 10.4 depicts the resulting changes of the parameter estimates and one parity equation between the LOLIMOT model for the nominal values and the RLS model for the faults.

Except for faults F3 and F4, which have the same physical effect on the condensate level in the heat exchanger, all faults show different patterns and are therefore isolable. As all considered faults lead to a deviation of the residual r_{ϑ_1}, this can be used for fault detection in normal operation, once the model (10.3.4) is identified. Then dynamic excitation of the three considered inputs $\dot{V}_1$, $\dot{m}_s$ and ϑ_{1i} can be started, e.g. one after another, to obtain parameter estimates in order to diagnose the faults, according to the symptoms given in Table 10.4. The fault detection can be further improved by adaptive thresholds and fuzzy-logic-based fault diagnosis, [10.3]. In this case two flow rates and two temperatures were required as measurements, assuming a constant steam pressure.

Table 10.4. Fault-symptom table for the steam/water heat exchanger. + increased; − decreased; 0 no effect (for sensor faults increased gains are considered)

faults	symptoms					
	parameter estimation					parity equation
	$K_{\dot{V}_1}$	$K_{\dot{m}_s}$	K_{ϑ_1}	T_{1o}	c_0	r_{ϑ_1}
F1 leak	−	+	−	+	+	+
F2 cond. valve stuck open	+	0	0	−	0	+
F3 cond. valve stuck closed	+	−	+	−	+	+
F4 leak vacc. pump	+	−	+	−	+	+
F5 inert gas	+	−	+	−	0	−
F6 sensor ϑ_{1i}	0	0	−	0	0	−
F7 sensor $\dot{V}_1$	−	+	0	+	0	+
F8 sensor ϑ_{1o}	+	+	+	0	0	−
F9 sensor $\dot{m}_s$	0	+	0	0	0	−

10.4 Conclusions

Fault detection by applying *static models* for heat exchangers with two different fluids is based on heat balance equations and requires therefore one or two flow rate sensors and four temperature sensors as the examples in Section 10.2.2 have shown. These models then usually do not allow a detailed fault detection because of the overall balance or the restricted possibility to extract only a characteristic quantity where several parameters are lumped together.

Fault detection is improved if *dynamic models* are used, because then more parameters can be estimated which change differently under the influence of faults and therefore improve the isolability. This holds especially if several variables, like two fluids and two temperatures can be dynamically changed, as Section 10.3.2 has shown. Also fewer sensors may be used as with overall static models.

Part IV

Fault-tolerant Systems

11

Fault-tolerant systems – a short introduction

The improvement of reliability can be increased by two different approaches, *perfectness or tolerance*, [11.4]. Perfectness refers to the idea of avoiding faults and failures by means of an improved mechanical or electrical design. This includes the continued technical advancement of all components that increase the operational life. During operation the intactness of the component must be maintained by regular maintenance and replacement of wearing parts. Methods that facilitate fault detection at an early stage allows one to replace the regular maintenance schedule with a maintenance-on-demand scheme.

Tolerance describes the notion of trying to contain the consequences of faults and failures such that the components remain functional. This can be reached by the principle of *fault tolerance*. Herewith, faults are compensated in such a way that they do not lead to system failures. The most obvious way to reach this goal is *redundancy* in components, units or subsystems called modules. However, the overall systems then become more complex and costly. In the following, various types of fault-tolerant methods are reviewed briefly, for more details see [11.3] and [11.8].

Fault-tolerance methods generally use *redundancy*. This means that in addition to the considered module, one or more modules are connected, usually in parallel. These redundant modules are either *identical* or *diverse*. Such redundant schemes can be designed for hardware, software, information processing, and mechanical and electrical components like sensors, actuators, microcomputers, buses, power supplies, etc.

11.1 Basic redundant structures

There exist mainly two basic approaches for fault tolerance: static redundancy and dynamic redundancy. The corresponding configurations are first considered for *electronic hardware* and then for other components. Figure 11.1a) shows a scheme for *static redundancy*. It uses three or more parallel modules that have the same input signal and are all active. Their outputs are connected to a voter, which compares these signals and decides by majority which signal value is the correct one. If a triple

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_11, © Springer-Verlag Berlin Heidelberg 2011

modular-redundant system is applied, and the fault in one of the modules generates a wrong output, this faulty module is masked (i.e. not taken into account) by the two-out-of-three voting. Hence, a single faulty module is tolerated without any effort for specific fault detection and n redundant modules can tolerate $(n-1)/2$ faults (n odd).

To improve the fault tolerance also the voter can be made redundant, [11.8]. Disadvantages of static redundancy are high costs, more power consumption and extra weight. Furthermore, it cannot tolerate common-mode faults, which appear in all modules because of common fault sources.

Dynamic redundancy needs fewer modules at the cost of more information processing. A minimal configuration consists of two modules, Figure 11.1b) and c). One module is usually in operation and, if it fails, the standby or back-up unit takes over. This requires fault detection to observe if the operating modules become faulty. Simple fault-detection methods only use the output signal for, e.g. consistency checking (range of the signal), comparison with redundant modules or use of information redundancy in computers like parity checking or watchdog timers. After fault detection, it is the task of the reconfiguration to switch to the standby module and to remove the faulty one.

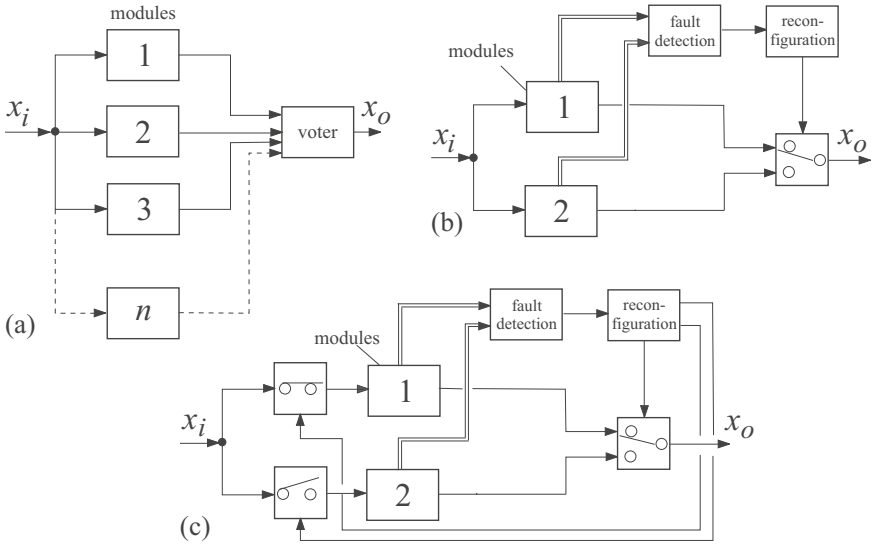


Fig. 11.1. Fault-tolerant schemes: a) static redundancy: multiple-redundant modules with majority voting and fault masking, m out of n systems (all modules are active); b) dynamic redundancy: standby module that is continuously active, "hot standby"; c) dynamic redundancy: standby module that is inactive, "cold standby"

In the arrangement of Figure 11.1b), the standby module is continuously operating, called *hot standby*. Then, the transfer time is small at the cost of operational aging (wear-out) of the standby module.

Dynamic redundancy where the standby module is not in operation and does not wear is shown in Figure 11.1c) and is called *cold standby*. This arrangement needs two more switches at the input and more transfer time due to a start-up procedure. For both schemes, the performance of the fault detection is essential.

Dynamic redundancy can be extended to two and more standby modules, thus tolerating two or more faults. Combinations of static and dynamic redundancy lead to *hybrid redundant schemes* to avoid the disadvantages of both types on cost of higher complexity, [11.8].

Table 11.1 gives a brief summary of static and dynamic redundancy as applicable to electronic hardware, software, mechanical and electrical systems and mechatronic systems. More details are described in [11.3], [11.5], [11.8].

11.2 Degradation steps

Mainly because of costs, space and weight, a suitable compromise between the degree of fault tolerance and the number of redundant modules has to be found. In contrast to fly-by-wire systems, only one or two failures can be tolerated for hazardous cases, for industrial and traffic systems, mainly because a safe state can be reached easier and faster. This means that not all components need very stringent fault-tolerance requirements. The following steps of degradation are distinguished:

- *fail-operational* (FO): one failure is tolerated, i.e. the module stays operational after one failure. This is required if no safe state exists immediately after the component fails
- *fail-safe* (FS): after one (or several) failure(s), the module directly possesses a safe state (passive fail-safe, without external power) or is brought to a safe state by a special action (active fail-safe, with external power)
- *fail-silent* (FSIL): after one (or several) failure(s), the module is quiet externally, i.e. stays passive by switching off and therefore does not influence other components incorrectly
- *fail* (F): permanent interruption of the module's ability to perform a required function.

For, e.g. vehicles, it is proposed to subdivide FO into "long time" and "short time," in order to reach a safe state dependent on the kind of failure. Considering these degradation steps for various components, one has to check first if a safe state exists. For automobiles, (usually) a safe state is standstill (or low speed) at a non-hazardous place. For components of automobiles, a fail-safe status is (usually) a mechanical back-up (i.e. a mechanical or hydraulic linkage) for direct manipulation by the driver. Passive fail-safe is then reached, e.g. after failure of electronics if the vehicle comes to a stop independently of the electronics, e.g. by a closing spring in the throttle or by actions of the driver via mechanical backup. However, if no mechanical back-up exists after failure of electronics, only an action by other electronics (switch to a still operating module) can bring the vehicle (in motion) to a safe state, i.e. to reach a stop through active fail-safe. This requires the availability of electrical power.

Table 11.1. Static and dynamic redundancy with multiple modules (examples)

Systems	Fault-tolerance methods							
	Static redundancy				Dynamic redundancy			
	Type	Modules	Selection	Fault masking	Type	Modules	Fault detection	Reconfiguration
Electronic hardware	Static redundancy n-out-of-m	Identical $n \geq 3$	Majority logic voting	Yes	Dynamic redundancy: hot standby, cold standby	Identical $n \geq 2$	Simple output checks	Switch to standby module: output, input and output
Software	Multiple version programming	Repeated running	Majority logic voting	Yes	Recovering blocks (like cold standby)	Diverse program modules	Acceptance tests with outputs	Switch to standby program and back to recovery point
		Diverse programs $n \geq 3$	Majority logic voting	Yes				
Mechanical systems	Static redundancy	Identical $n \geq 2$	Physical compatibility	No	Dynamic redundancy: cold standby	Identical or diverse $n \geq 2$	Simple output checks	Switch to standby unit
Electrical systems	Static redundancy	Identical $n \geq 2$	Physical compatibility	No	Dynamic redundancy: cold standby	Identical or diverse $n \geq 2$	Simple output checks	Switch to standby unit
Mechatronic systems	–	–	–	–	Dynamic redundancy: cold standby	Identical or diverse $n \geq 2$	Sophisticated: inputs & outputs	Switch to standby unit

Generally, a *graceful degradation* is envisaged, where less critical functions are dropped to maintain the more critical functions available, using priorities, [11.2]. Table 11.2 shows degradation steps to fail-operational (FO) and fail (F) for different redundant structures of electronic hardware. As the fail-safe status depends on the considered system and the kind of components, it is not considered here.

Table 11.2. Fail behavior of electronic hardware for different redundant structures. FO: fail-operational; F: fail; FS: fail-safe not considered

		Static redundancy		Dynamic redundancy		
Structure	Number of elements	Tolerated faults	Fail behavior	Tolerated failures	Fail behavior	Discrepancy detection
Duplex	2	0	F	0	F	2 comparators
				1	FO-F	fault detection
Triplex	3	1	FO-F	2	FO-FO-F	fault detection
Quadruplex	4	1	FO-F	3	FO-FO-FO-F	fault detection
Duo-duplex	4	1	FO-F	–	–	–

For flight-control computers, usually a triplex structure with dynamic redundancy (hot standby) is used, which leads to FO-FO-FS, such that two failures are tolerated and a third one allows the pilot to operate manually, [11.1], [11.6], [11.7]. If the fault tolerance system has to cover only one fault to stay fail-operational (FO-F), a triplex system with static redundancy or a duplex system with dynamic redundancy is appropriate. If fail-safe can be reached after one failure (FS), a duplex system with two comparators is sufficient. However, if one fault has to be tolerated to continue fail-operational and after a next fault it is possible to switch to a fail-safe (FO-FS), either a triplex system with static redundancy or a duo-duplex system may be used, see [11.7]. The duo-duplex system has the advantages of simpler failure detection and modularity.

12

Examples of fault-tolerant systems

High-integrity systems require a comprehensive overall fault tolerance by fault-tolerant components and an automatic fault management system. This means first the design and realization of redundant components which have the lowest reliability and are safety relevant. In automatically controlled systems there are, for example sensors, actuators, computers, communication (bus) systems, control and operational software and process parts, like electrical drives, tube lines, pumps or heat exchangers. Components with multiple redundancy are known for aircraft, space, train and nuclear power systems. Other technical processes with redundancy are for, example lifts (multiple ropes and brakes) or multiple pumps for steam boilers, see [12.2].

Some already-produced fault-tolerant systems and some prototype realizations are presented in this chapter. The next section begins with a general description of a fault-tolerant control system. This is followed by examples for fault-tolerant electrical drives, actuators and sensors, compare [12.40].

12.1 A fault-tolerant control system

This section describes the tasks of a fault-tolerant control system with automatic fault management taking components of control systems into account. It consists of fault-tolerant actuators, sensors and controllers and is represented in Figure 12.1. This scheme is considered first, because sophisticated fault-tolerant systems frequently operate automatically, either in open or closed loop. The considered fault-tolerant system may consist of:

(i) *Fault-tolerant actuators*:
- redundant identical or diverse actuators
- actuators with inherent fault tolerance
- actuator reconfiguration module.

(ii) *Fault-tolerant sensors*:
- redundant identical or diverse sensors
- sensors with inherent fault tolerance

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_12, © Springer-Verlag Berlin Heidelberg 2011

- virtual sensors based on analytical redundancy
- sensor reconfiguration module.

(iii) *Active fault-tolerant controllers*:

- redundant identical or diverse controller hardware
- redundant diverse controller software
- different controller structure and parameters
  - pre-designed for a priori-known faults
  - redesigned or adaptive after fault detection.

(iv) *Fault-detection module*:

- normal closed-loop operating signals are used to detect and isolate faults in the components (parity equations, observers, parameter estimation)
- test signals are introduced either periodically or on request to improve fault detection and, if required, fault diagnosis ("active fault detection")
- indication of the degree of impairment and degree of safety criticality.

(v) *Fault-management module*:

- decisions based on fault detection with an indication of the degree of impairment and effect on safety of the components
- reconfiguration strategies with
  - hard or soft reconfiguration
  - change of operating conditions (setpoints, process performance)
  - closed-loop or open-loop (feedforward) operation.

The scheme in Figure 12.1 is an example with two manipulated and two controlled variables. If the normal *actuator 1 fails*, e.g. by getting stuck, actuator 2 replaces its functions. This can be a second actuator of the same type or another actuator with a similar manipulation effect on the control variable. For example, certain additional control surfaces for aircraft can be used as other, redundant actuators, e.g. ailerons or rudders.

In the case that the fault-detection system detects that *sensor 1* has a fault or even fails totally a second sensor of the same type is switched or some analytical redundancy with other sensors is used to generate a virtual sensor output 2. (Examples are the electrical throttle with a double potentiometer, [12.19] or the model-based calculation of the yaw rate of automobiles from the lateral acceleration and wheel speed sensors, [12.20] or a horizontal and vertical gyro for the bank angle of aircraft, [12.44].)

Depending on the reconfigured actuators or sensors, the controller structure and/or controller parameters of the *fault-tolerant controller* have also to be reconfigured.

The structure of Figure 12.1 also holds for *faults in the process* itself. If actuator 2 or sensor 2 can be used to maintain the operation, the reconfiguration just selects the actuator–sensor configuration, adjusts the controller 2 and the reference variable w_2 accordingly. An example is a fluid 1 / fluid 2 heat exchanger: The outlet temperature of fluid 1 can be manipulated by changing the flow of fluid 2 instead of the temperature of fluid 2, if the plant allows.

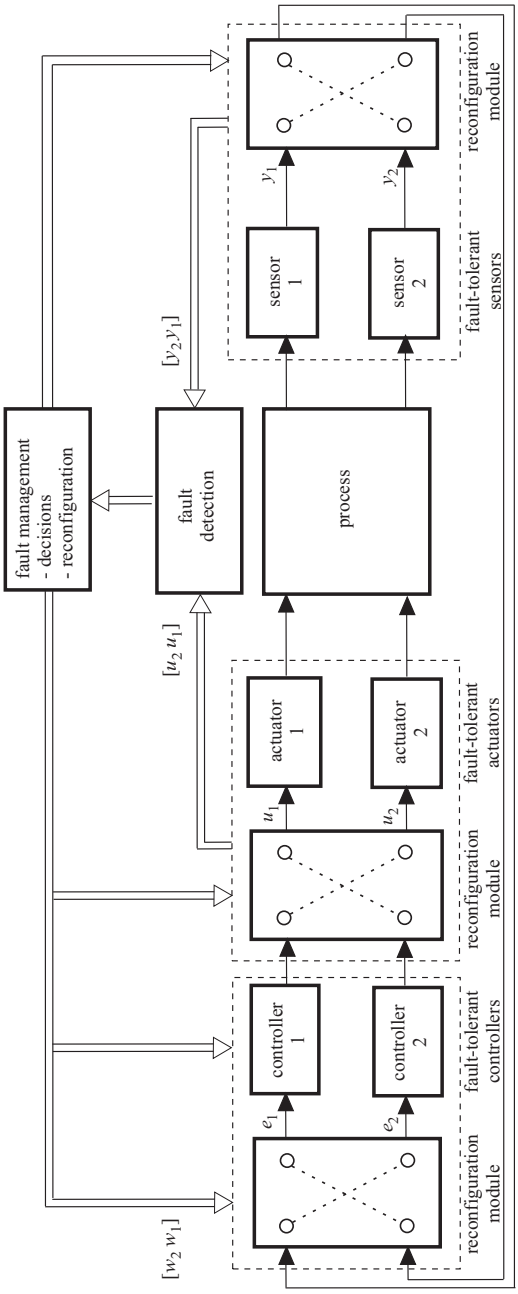


Fig. 12.1. Fault-tolerant control system with automatic fault management, shown for 2 actuators, 2 sensors and 2 controllers

The *controller structure and parameters* have to be adapted to the new process behavior. If the transfer behavior of the reconfigured actuator–process–sensor system is known in advance, preprogrammed controllers have just to be switched. If the behavior is not known, self-tuning or adaptive control algorithms could be used. However, this adaptation must be supervised and properly excited with perturbation signals, see [12.21], which can be a problem, if a very fast recovery is required.

The task of the *fault-detection module* is to detect faults in all components, like actuators, sensors, controllers and the process as early as possible. A diagnostic capability is not necessarily required, because it is mostly enough information for the reconfiguration to know if the actuator or sensor has failed, independent of the causes.

It should be mentioned that in the case of *closed-loop control* fault detection must be made under closed-loop conditions with all the problems discussed in [12.19], Chapter 12. Because of the danger of a reconfigured replacement controller not functioning as expected with a replacement sensor, it is sometimes better not to reconfigure an alternative closed-loop control, but to apply a *feedforward control* without an output sensor substitute. This may result in a loss of control performance, but instability is avoided. This is, for example, used in engine control. In the case that the oxygen sensor (λ-sensor) fails, a stoichiometric air/fuel ratio is maintained, based on air-flow measurement and the setpoint of injected fuel mass.

Faults in the *controller hardware or software* can be detected as described in [12.19], Chapter 12. Then new controllers as described above are applied or feedforward control is used.

In the following sections some examples for fault-tolerant drives, actuators and sensors are described, which are partially based on [12.40].

The discussion on automatic fault management shows that there are many different possibilities. Therefore it is difficult to treat applicable methods generally and it is recommended to consider concrete cases.

An experimental investigation of fault detection in closed loop and reconfiguration to a redundant sensor is shown for the electrical throttle valve actuator in [12.45] and [12.22].

12.2 Fault-tolerant electrical drives

Fault statistics of AC motors indicate that about 51% are due to bearing faults, 16% to stator windings, 16% to external equipment, 5% to bar and ring ruptures, 2% to shaft and couplings, and 10% to others, [12.19], [12.51], [12.53]. The faults are mostly caused by overload, overheating and missing lubrication. With regard to the 58% mechanical and also > 16% electrical faults a redundant, second motor is a first choice. Stator winding faults can be made fault tolerant by multi-phase designs. Both cases will be considered.

12.2.1 A fault-tolerant duplex AC motor

The construction of a redundant electromotor drive with two motors can be arranged either as a parallel or a serial structure, Figure 12.2. For a basic investigation two

standard inductance motors are used in a laboratory set-up, which are coupled by electro-mechanical clutches and drive a synchronous generator as a load. The parallel structure uses two belt-gears and two clutches, see Figure 12.3, whereas the serial structure only needs one clutch, [12.48]. The *parallel structure* allows one to switch off one faulty motor totally. A belt drive was chosen to ease a flexible laboratory arrangement. It could also be a toothed gear with two input shafts and one output shaft. The price for the parallel configuration is one gear and two clutches, which of course deteriorate the reliability. The *serial structure* is much simpler. If motor 2 is used the active drive, motor 1 and clutch can be cold standby and do not wear. However, redundance of motor 1 is only possible for electrical faults of motor 2 and not for breakdown because of mechanical faults (bearings, ruptures). Use of motor 1 as the active drive requires that motor 2 turns (without torque generation) and thus the mechanics are not completely cold standby, but only the electrical part. But motor 1 can be switched off totally, also in the case of a mechanical breakdown. Both configurations allow one to compensate one fault in one of the motors: F0-F. A special solution for a parallel configuration results if the two motors are coupled with a differential gear, [12.54]. However, this scheme needs two brakes, see also [12.38].

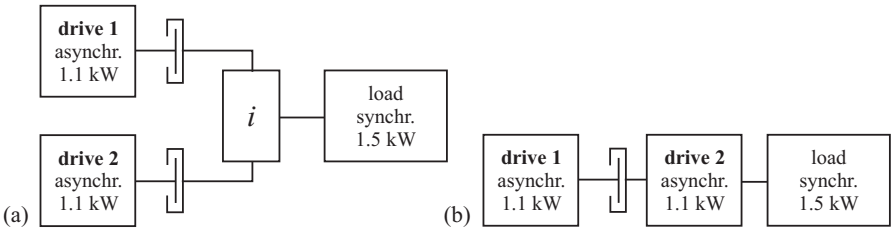


Fig. 12.2. Redundant AC-motor drive: a) parallel structure (2 electro-mechanical clutches, 1 gear); b) serial structure (1 electro-mechanical clutch)

The reconfiguration after fault detection will now be shown for the parallel structure. It is assumed that the active motor 1 is, after the detection of a fault, switched off and the cold-standby motor 2 is switched on, compare Figure 12.3. Herewith, clutch 1 has to be opened, clutch 2 closed and motor 2 started. The goal of this reconfiguration is to avoid a larger torque and speed reduction of the driven load. As the kind of reconfiguration depends on the detected fault, the single faults are given a fault measure

$$F_{tot} = \sum_{i=0}^{n} (g_i F_i); \quad F_i \in [0,1]; g_i \in \{0,1\} \tag{12.2.1}$$

For $F_{tot} < 1$ a soft-clutch switching and for $F_{tot} \geq 1$ a hard-clutch switching is triggered.

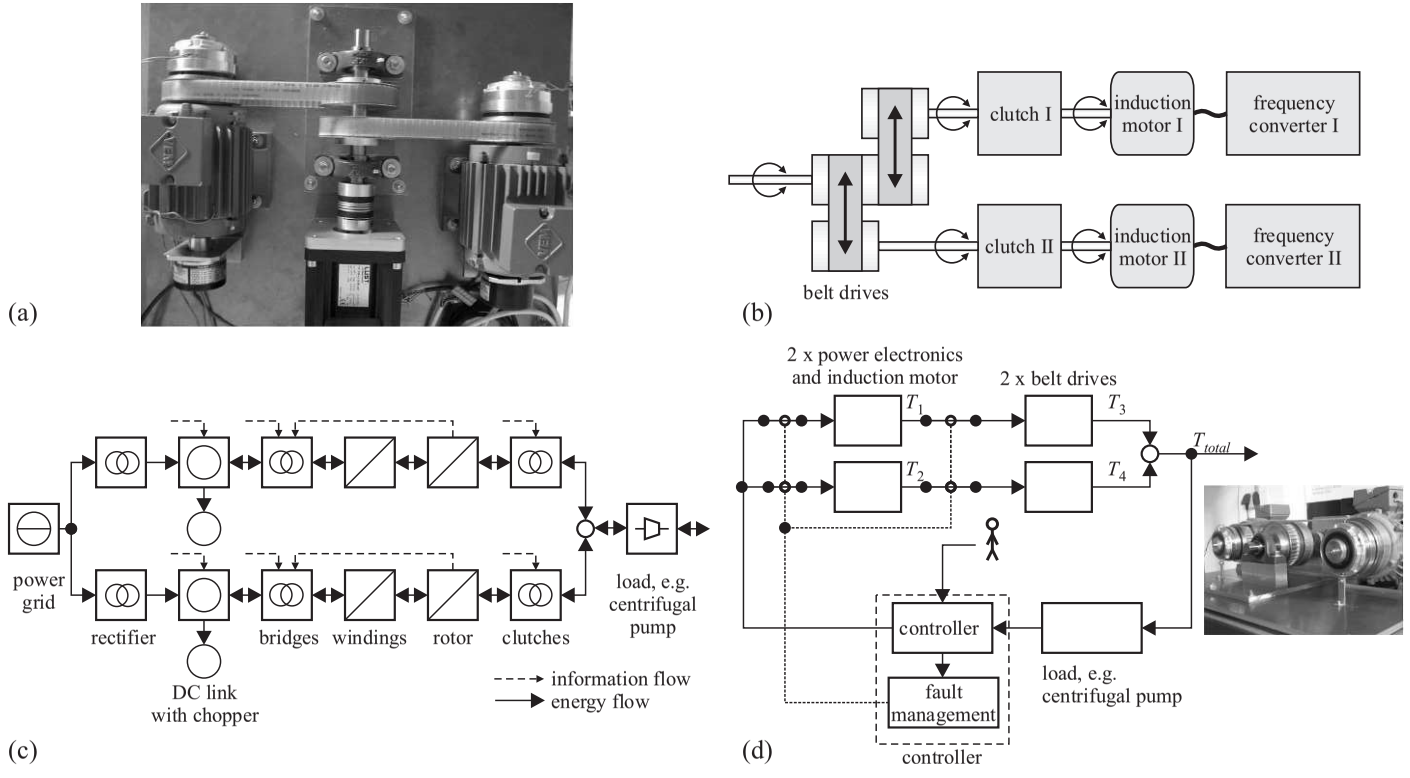


Fig. 12.3. Duplex AC motor drive: 2 belt drives with each one clutch on motor side, drive and emulated load (permanent magnetic synchronous motor): a) picture; b) schematic; c) energy flow scheme; d) signal flow diagram

a) Soft switching over: slow reconfiguration

If for a small fault it is still possible to operate motor 1, the switching over is performed softly. The stationary motor 2 is accelerated to the reference speed of motor 1. Then the opened clutch is closed stepwise at time t_1 by the PWM-command input $U_{K-2}(t)$ until a touch point. $U_{K-2}(t)$ then increases rampwise during $t_2 - t_1 = 0.5\,\mathrm{s}$ until t_2, see Figure 12.4. The clutch of the faulty motor is opened at t_2 with $U_{K-1}(t)$ in the opposite way and the motor is switched off. Figure 12.5 shows during an acceleration phase that a fault is detected at $t = 0.2\,\mathrm{s}$. The soft switching over is started and hardly any change of the load speed can be observed.

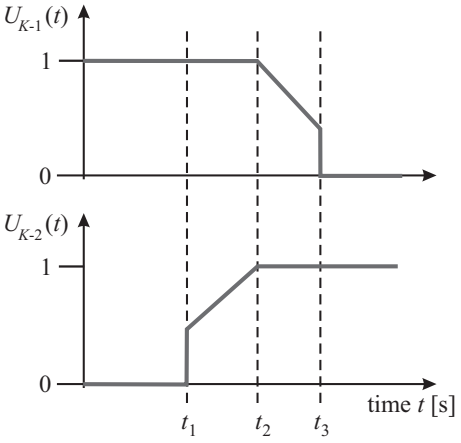


Fig. 12.4. Soft-clutch switching over

b) Hard switching over: fast reconfiguration

A severe fault, like in the power electronics, resulting in a large fault measure F_{tot} requires a fast reconfiguration. Then, also the electromagnetic clutches have to react fast. Therefore they are overexcited, resulting in a reduction of closing time from $T_{K-norm} = 130\,\mathrm{ms}$ to $T_K = 32\,\mathrm{ms}$. After fault detection at time t_1 the second motor is immediately started and accelerated to the reference speed. Based on the clutch time T_K and the acceleration $\dot{\omega}_2$ a lead speed difference

$$\Delta\omega_l(t) = T_K\ \dot{\omega}_2(t) \tag{12.2.2}$$

is calculated. Then, the second clutch is closed hardly at time t_2 if the clutch closing algorithm

$$\begin{aligned} U_{K-2}(t) &= \frac{1}{2} + \frac{1}{2}\ sgn\left[\omega_2(t) + \Delta\omega_l(t) - \omega_{load}(t)\right] \\ U_{K-2}(t) &\in \{0, 1\} \end{aligned} \tag{12.2.3}$$

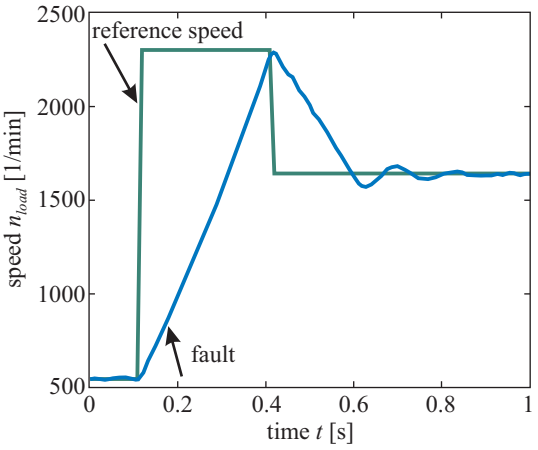


Fig. 12.5. Course of load speed for soft switching over at $t_1 = 0.2\,\mathrm{s}$

reaches the value 1 and the defective motor is switched off at t_3, see Figure 12.6. Figure 12.7 shows that with this reconfiguration only a small speed reduction arises. If the clutches would immediately be switched over, the speed reduction would be much larger, as also to be seen in Figure 12.7.

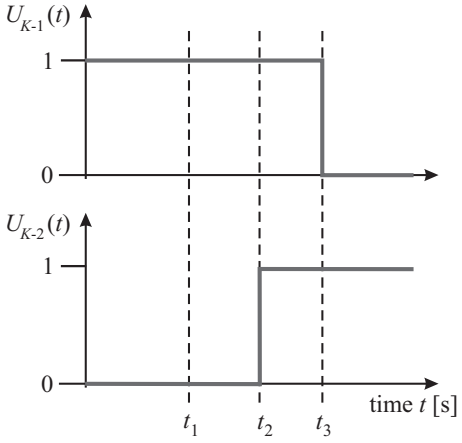


Fig. 12.6. Hard-clutch switching over

12.2.2 Fault-tolerant frequency converter

The power electronics for frequency-controlled AC drives have usually one inverter leg per motor phase. Upon the loss of one leg, the three-phase motor becomes single-phased and cannot generate a rotating magnetic field any longer. This is rather crit-

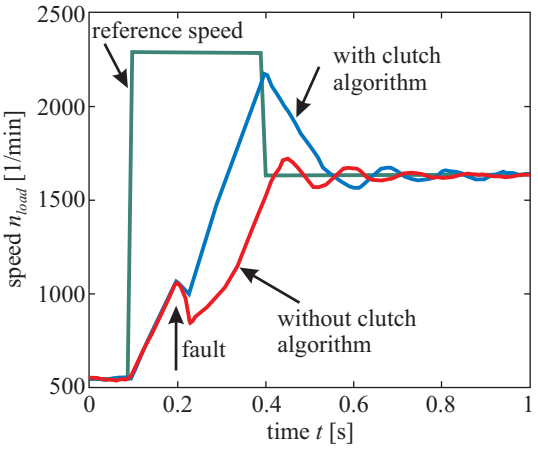


Fig. 12.7. Course of load speed for hard switching over at $t_1 = 0.2\,\mathrm{s}$ with and without clutch-control algorithm (12.2.3)

ical, because according to [12.24], the most common power inverter fault is in fact the loss of one leg, i.e. at least one permanently open switch. To overcome this fault situation, one can control each phase separately with a full H-bridge. If each phase of the motor is connected to a full H-bridge, the three-phase motor, upon the loss of one inverter leg or phase, can be operated as a two-phase motor and can still produce a rotating magnetic field. The major disadvantage is the fact that for separate H-bridges, each winding needs two wires to establish the connection to the power converter, making the wiring more expensive, especially for machines which are placed far away from the power electronics. Upon the loss of a phase, the voltage of the DC link must be increased to maintain the power rating, [12.28]. In order to be able to increase the DC-link voltage, the rectifier part of Figure 12.8 must be retrofitted with active switches instead of the passive switches, i.e. diodes. Furthermore, during normal operation, the voltage level provided by the power grid is not fully utilized, thus performance is degraded. The combination of a standard three-phase PMSM (permanent-magnet synchronous motor) with an inverter with full H-bridges has been presented in [12.15], [12.26], [12.29].

The system is now also FO-FSIL with respect to inverter or motor phase faults. Upon the first fault at a motor phase or inverter leg, the motor becomes two-phased. Upon the second loss, all phases are disconnected and the motor is at least not generating any torque due to induced currents, so it is fail-silent with respect to impairments of the load. There is also a degradation step: Upon the loss of one phase, the torque produced becomes unevenly distributed along the circumference of the stator. An overview of many other fault-tolerant frequency inverter topologies is presented in [12.38].

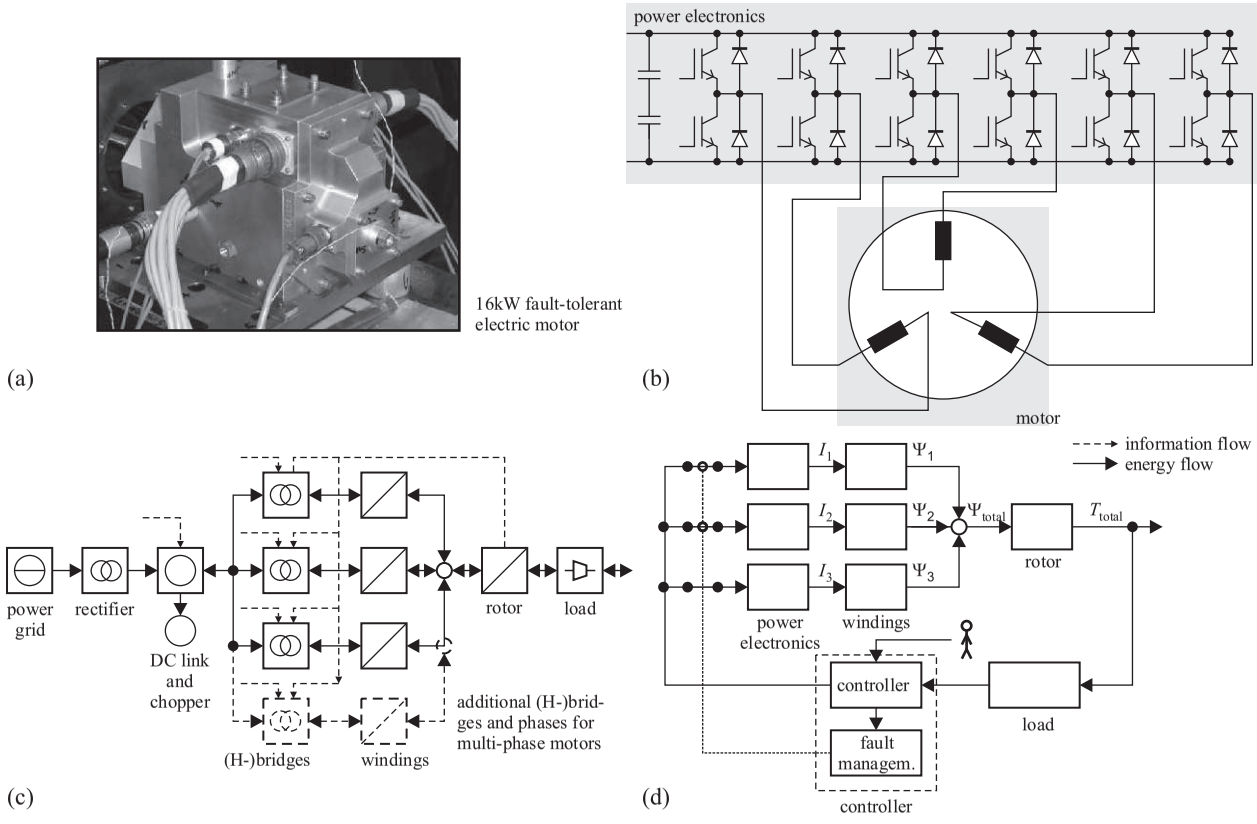


Fig. 12.8. Fault-tolerant electrical drive with fault-tolerant frequency converter: a) schematic; b) energy flow scheme; c) block diagram, see [12.1], [12.15], [12.25], [12.26], [12.29]

12.2.3 Multi-phase motors

An alternative to using full H-bridges is to design so called *multi-phase motors*, that have more than three phases. Although the first designs date back to the late 1960s, [12.28], they have only recently been in the focus of research, as they are well suited for applications in the area of ship, locomotive, and electric/hybrid vehicle propulsion, see, e.g. [12.28], and the more electric aircraft, [12.13], [12.49], an initiative gathering momentum in the late 1990s with the aim to control aircraft subsystems with electrically actuated drives in place of mechanical, hydraulic or pneumatic means.

One reason for the introduction of multi-phase motors in applications demanding the highest power is that semiconductors are not yet capable of switching the high currents that traditional three-phase motors would need to satisfy such high power demands. By increasing the number of phases, the power-per-switch can be limited to values that can be borne by the semiconductors. Also, multi-phase machines can easily sustain phase losses. Upon the loss of one phase, an n-phase motor becomes an $(n-1)$ phase motor ($n \geq 4$) and hence shows a smaller loss in the power-rating and the uniformity of the circumferential torque distribution as the number of phases n increases. Note that the stator windings now have no common starprint. Furthermore, multi-phase motors show improvements in the noise characteristics and their torque production can easily be enhanced by the injection of higher-frequency harmonics.

The design of a four-phase fault-tolerant PMSM aircraft actuator is shown in [12.1]. A five-phase permanent magnet motor has been realized by [12.5] and has been investigated experimentally for post-fault operation. Depending on the number of windings, the topology of the windings and the allowable loss in torque, the system can sustain one or more faults, that means the system is *at least* FO-FSIL with respect to phase faults.

12.3 Fault-tolerant actuators

Actuators generally consist of different parts: input transformer, actuation converter, actuation transformer and actuation element (e.g. a set of DC amplifier, DC motor, gear and valve), as shown in Figure 12.9a). The actuation converter converts one form of energy (e.g. electrical or pneumatic) into another form (e.g. mechanical or hydraulic). Available measurements are frequently the input signal U_i, the manipulated variable U_1 and an intermediate signal U_3.

Fault-tolerant actuators can be designed by using multiple complete actuators in parallel, either with static redundancy or dynamic redundancy with cold or hot standby (Figure 11.1). One example of static redundancy is hydraulic actuators for fly-by-wire aircraft where at least two independent actuators operate with two independent hydraulic energy circuits.

Another possibility is to limit the redundancy to parts of the actuator that have the lowest reliability. Figure 12.9b) shows a scheme where the actuation converter (motor) is split into separate parallel parts. Examples with static redundancy are two

servo-valves for hydraulic actuators, [12.4], [12.3] [12.43] or multiple windings of an electrical motor (including power electronics), [12.25], see also [12.19]. Within electromotor-driven throttles for SI engines, only the slider is doubled to make the potentiometer position sensor static-redundant, see [12.18].

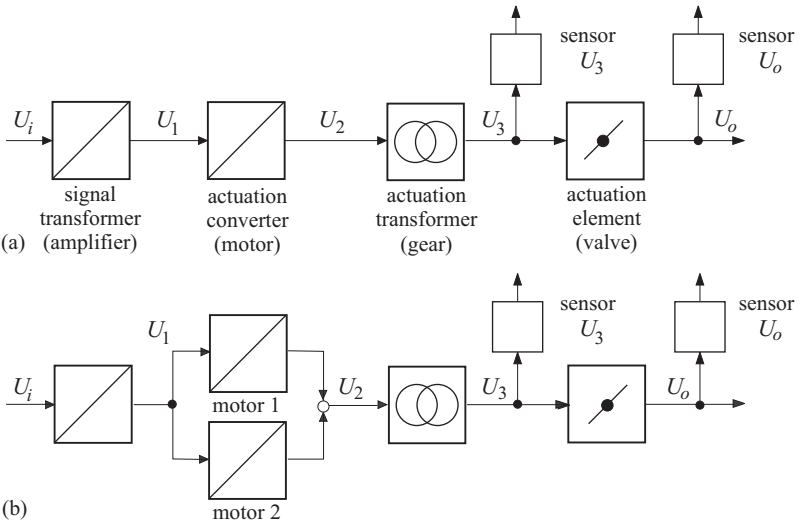


Fig. 12.9. Fault-tolerant actuator: a) common actuator; b) actuator with duplex drive

One example for dynamic redundancy with cold standby is the cabin pressure flap actuator in aircraft, where two independent DC motors exist and act on one planetary gear, [12.33], see [12.19].

As cost and weight generally are higher than for sensors, actuators with fail-operational duplex configuration are to be preferred. Then, either static-redundant structures, where both parts operate continuously, Figure 11.1a), or dynamic redundant structures with hot standby, Figure 11.1b), or cold standby, Figure 11.1c), can be chosen. For dynamic redundancy fault-detection methods of the actuator parts are required, [12.47]. One goal should always be that the faulty part of the actuator fails silently, i.e. has no influence on the redundant parts.

12.3.1 Fault-tolerant hydraulic actuators

Research into fault-tolerant hydraulic systems has been sparked mainly by aeronautical applications, since hydraulic actuators are used in many actuation tasks in aircraft. They will also prevail for the control of primary flight control surfaces in the future. However, the secondary flight control surfaces might be replaced by electro-hydraulic actuators, [12.9], [12.10], [12.13], [12.23], [12.49], [12.52].

In [12.13], electro-hydraulic and electro-mechanical actuation are compared with respect to their advantages and disadvantages for aeronautical use. Here, the electro-

hydraulic actuator (EHA) is recommended due to fact that the *reliability of EHA components* has been *studied well* and that both hot-standby and cold-standby configurations have been developed and validated. The major disadvantages are the *maintenance* necessary for hydraulic components and, in the case of distributed pressure supplies, the *cost and weight*. The negative effects of a distributed pressure supply can however obviously be reduced by the use of central (regulated) pressure supplies. Yet, central pressure supplies do also exhibit severe disadvantages: The *infrastructure* (i.e. the piping) *is heavy and inflexible* and in the case of leakages *corrosive fluids might spill*, see [12.49].

On the contrary, the electro-mechanical actuator (EMA) can be mainly seen as a stand-by actuator as in aeronautical applications. This actuation principle is currently limited to the actuation of secondary flight surfaces. Here, the *significant maintenance cost reduction* due to the *reduction of wearing parts*, such as seals, are very favorable, [12.6], [12.14]. The electro-mechanical drive is heavily impaired by the fact that a linear motion can in most cases only be generated by a matching gear, thus fostering *flutter concerns* due to the *free-play of the mechanical transmission and the jam susceptibility*. Upon shorts, there is also always the *risk of fire*, see [12.49].

a) Fault-tolerant dual valve and dual piston

Figure 12.10 shows the electro-hydraulic rudder actuator of the Eurofighter, see, e.g. [12.11], [12.27]. The actuator consists of a proportional-acting valve, which is driven by four separate solenoid units. Each one is supplied by its own power electronics. All solenoids are acting on the same valve spool. Two control edges each supply one cylinder chamber with hydraulic fluid. The cylinder has four chambers with identical active piston areas. As can be seen from the block diagram in Figure 12.10d), the valve spool and the piston rod are the most critical components. If the valve spool or the piston rod jams, the system cannot operate any longer. Furthermore, upon a jam of the valve spool, the piston rod (and the attached control surface) may run away. The system can sustain leakages and/or pressure losses in one of the two hydraulic circuits. It can also sustain internal leakages in one of the two hydraulic circuits without losing the characteristic stiffness of hydraulic systems. As long as neither the valve spool nor the piston rod are affected, the system is FO-F.

b) Fault-tolerant single piston with double valve

A different prototypical realization of a hydraulic actuator is shown in Figure 12.11. Here, only the valve has been doubled since it has been found by a detailed statistical analysis of maintenance records that the valve alone makes up roughly 51% of all faults at hydraulic servo axes, [12.36], [12.41]. As can be seen from the block diagram in Figure 12.11d), valve spool faults are now less critical as there are two independent valve spools in this design. Upon the jam of one valve spool, the other valve can take over the volume flow. Furthermore, the cylinder must no longer have equal active piston areas, which means that a differential cylinder can be used instead of a double-rod cylinder. This reduces the necessary installation space. Furthermore,

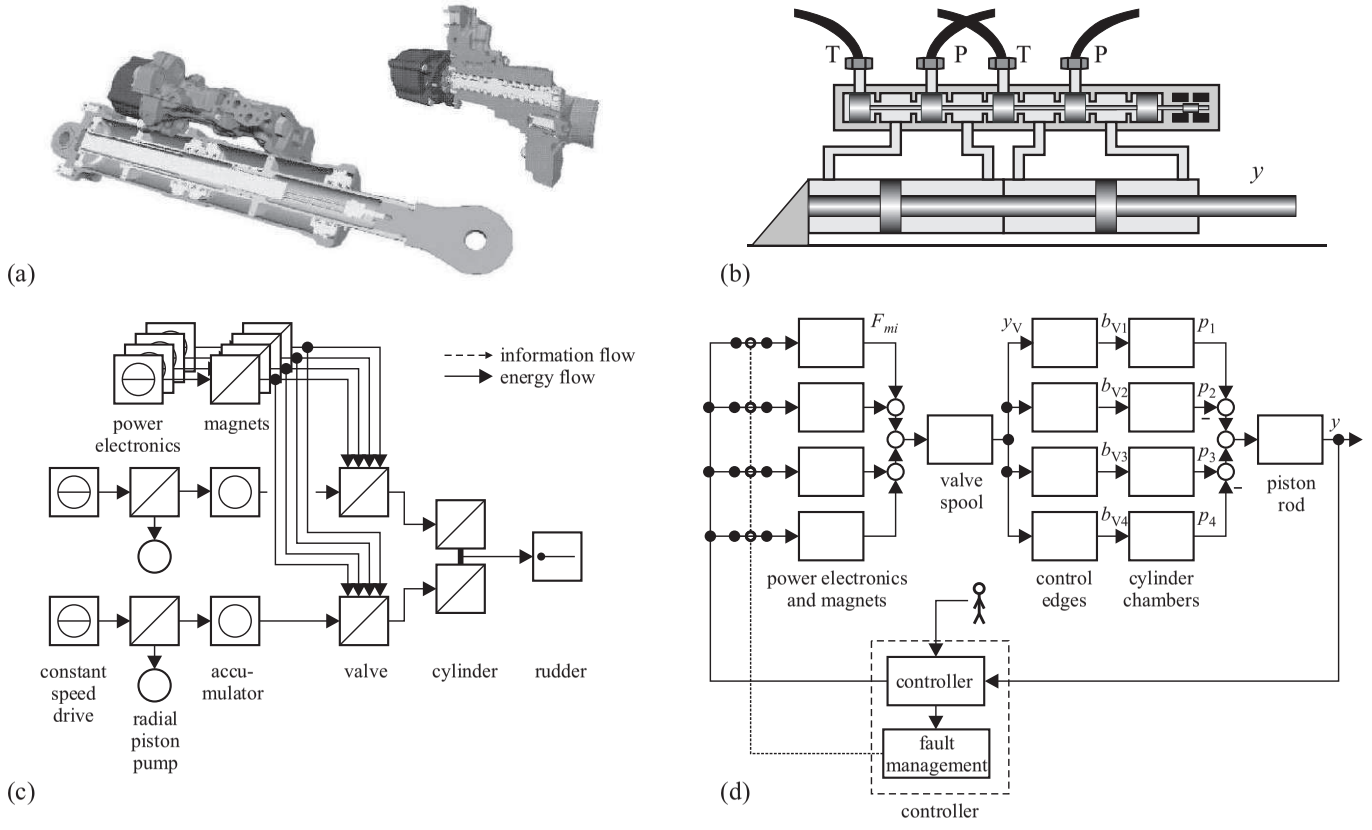


Fig. 12.10. Fault-tolerant electro-hydraulic rudder actuator for Eurofighter: a) system; b) schematic; c) energy flow scheme; d) block diagram, in d) redundancy by multiple onboard computers is not shown, see [12.11], [12.27]

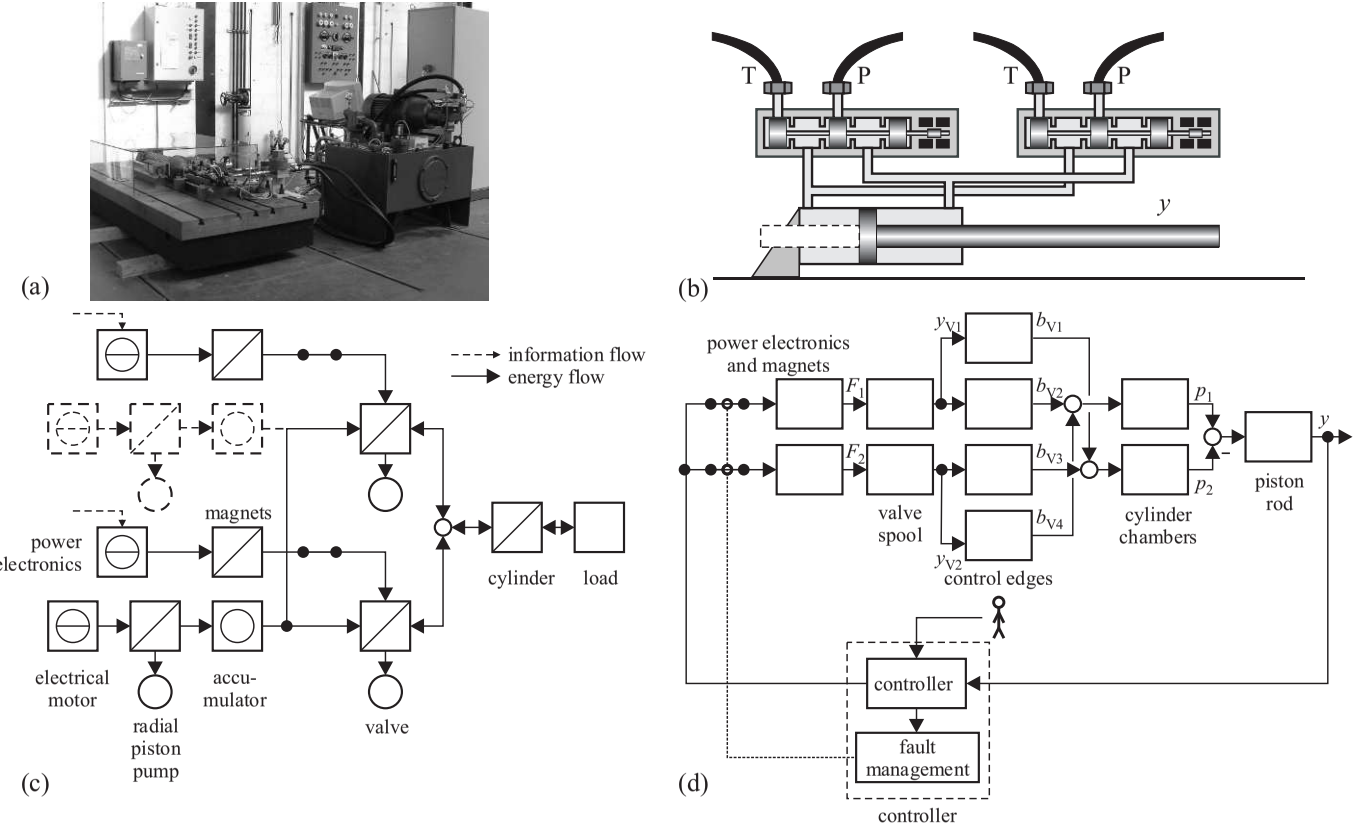


Fig. 12.11. Fault-tolerant electro-hydraulic servo axis: a) system; b) schematic; c) energy flow scheme; d) block diagram, see [12.41]

the entire servo axis is built from standard components, i.e. no design and construction of new, specialized components are necessary. Besides actuator fault tolerance, also sensor fault tolerance by means of analytical redundancy has been implemented. It is possible to operate the hydraulic servo axis in closed-loop position control even after the loss of the position sensor, see [12.37].

This setup can tolerate all valve faults. Even if one valve spool is blocked in a partly open position, the other valve can still be used in many situations to compensate this parasitic volume flow, [12.4], [12.3]. As the cylinder causes only 16% of all faults at hydraulic servo axes, it can be regarded as rather reliable. Furthermore, internal leakages only affect the stability and the stiffness of the drive, but do not endanger the further operation. As long as the piston rod is not affected and the spare valve is able to conduct any remaining parasitic flow of the faulty valve, the system is FO-F.

c) Fault-tolerant dual valve and pump

Another design with a dual-tandem ram, that is typical for aeronautical applications (see, e.g. [12.31], [12.42]), is shown in Figure 12.12. This figure shows an actuation system for the F/A-18 horizontal stabilizer, which is a secondary control surface. In this example, the hydraulic cylinder is doubled and directly supplied with hydraulic fluid by two fixed-displacement pumps that are driven by two brushless DC motors. Bypass valves allow the piston to move even if the motor axle, respectively pump jams. The big advantage of this setup is that in the case of removal of the component, no hydraulic connections must be loosened. Furthermore, in the case of a leakage of the hydraulic piping/components, only the hydraulic oil of the component specific hydraulic circuit will spill, leaving all other hydraulic circuits of the plane unaffected. Provided that the valves are still functional and can disconnect the piston chambers from the pump, the system is FO-FSIL.

Many other architectures for EHAs in aeronautical applications have been assessed in [12.50]. The big lead in X-by-wire functionalities and the accompanying use of fault-tolerant mechatronic components in aeronautical applications is typical and can be explained easily: Fly-by-wire functionalities for airplanes could be introduced with little increase in risk as control surfaces by themselves are redundant. Almost all maneuvers can be realized by different combinations of control surfaces.

12.3.2 Fault-tolerant DC actuator

Figure 12.13 gives an example of a fault-tolerant electro-mechanic component from the civilian aviation, see [12.33], [12.34]. It shows a *cabin outflow valve*, which is used to control the air pressure inside the fuselage. At the front of passenger planes, bleed air from the engines is injected into the cabin and at the rear, cabin outflow valves control the discharge of the pressurized air from the fuselage to the surrounding atmosphere.

There are typically two, four, or even more cabin outflow valves. Each valve has the same design as depicted in Figure 12.13: Two brushless DC motors with separate

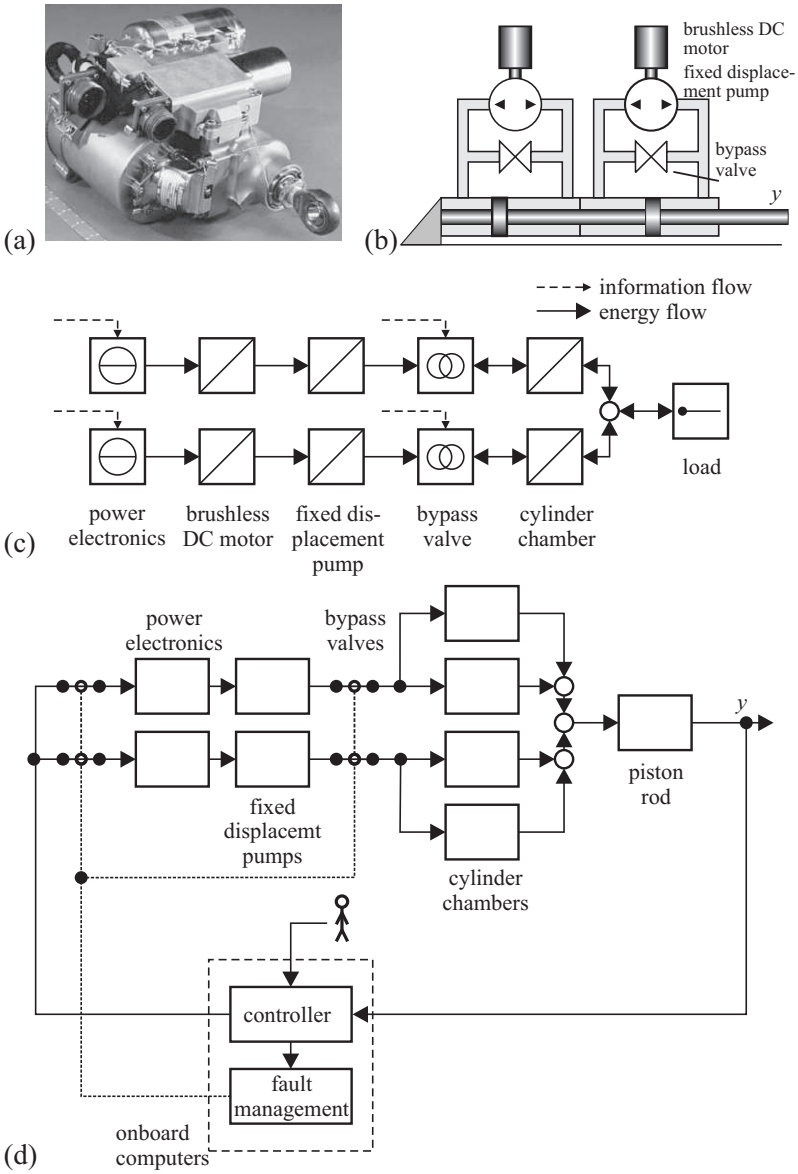


Fig. 12.12. Fault-tolerant electro-hydraulic servo axis: a) system; b) schematic; c) energy flow scheme; d) block diagram, (redundancy by multiple onboard computers is not shown) [12.31]

power electronics act on a common gear. Both power electronics are connected to different buses. The switch-over from one actuator to the other is initiated automatically. Both brushless DC motors operate in closed-loop cabin pressure control.

Fault detection, diagnosis and management is assigned to the pilot. Upon the loss of a brushless DC motor, the pilot can initiate a switch-over from the defect to the intact component. When both brushless DC motors fail, the cabin pressure control is then completely assigned to the pilot, who operates the valve directly via a third classical DC motor.

In this example, only the electronics and the electro-mechanical converter are made redundant, because the failure rate of electronic and electro-mechanical converters is much larger than that of mechanical parts. As part of the research project [12.32], [12.33], [12.35], fault detection and diagnosis methods have been developed, which allow one to detect and diagnose the following faults: Over-temperature, shorted winding, increased friction, offset faults of all sensors, stuck-at faults of the (binary) hall sensors. The system presents a setup with cold standby, as only one component is active at any time. It is FO-FO-F, i.e. fail operational after the first fault, also fail operational after the second fault and finally fails after the third fault at the electric drives.

12.4 Fault-tolerant sensors

12.4.1 Hardware sensor redundancy

Sensor systems with static redundancy are realized, for example, with a triplex system and a voter, Figure 12.14a). A configuration with dynamic redundancy needs at least two sensors and fault detection for each sensor, Figure 12.15b). Usually, only hot standby is feasible. Another less powerful possibility is plausibility checks for two sensors, also by using signal models (e.g. variance) to select the more plausible one, Figure 12.15c).

The fault detection can be performed by *self-tests*, e.g. by applying a known measurement value to the sensor. Another way uses *self-validating sensors*, [12.8], [12.17], where the sensor, transducer and a microprocessor form an integrated, decentralized unit with self-diagnostic capability. The self-diagnosis takes place within the sensor or transducer and uses several internal measurements, see also [12.30]. The output consists of the sensor's best estimate of the measurement and a validity status, like good, suspect, impaired, bad and critical.

12.4.2 Analytical sensor redundancy

As a simple example, a process with one input and one main output y_1 and an auxiliary output y_2 is considered, see Figure 12.15a). Assuming the process input signal u is not available but two output signals y_1 and y_2, which both depend on u, one of the signals, e.g. $\hat{y}_1$ can be reconstructed and used as a redundant signal if process models G_{M1} and G_{M2} are known and considerable disturbances do not appear (ideal cases).

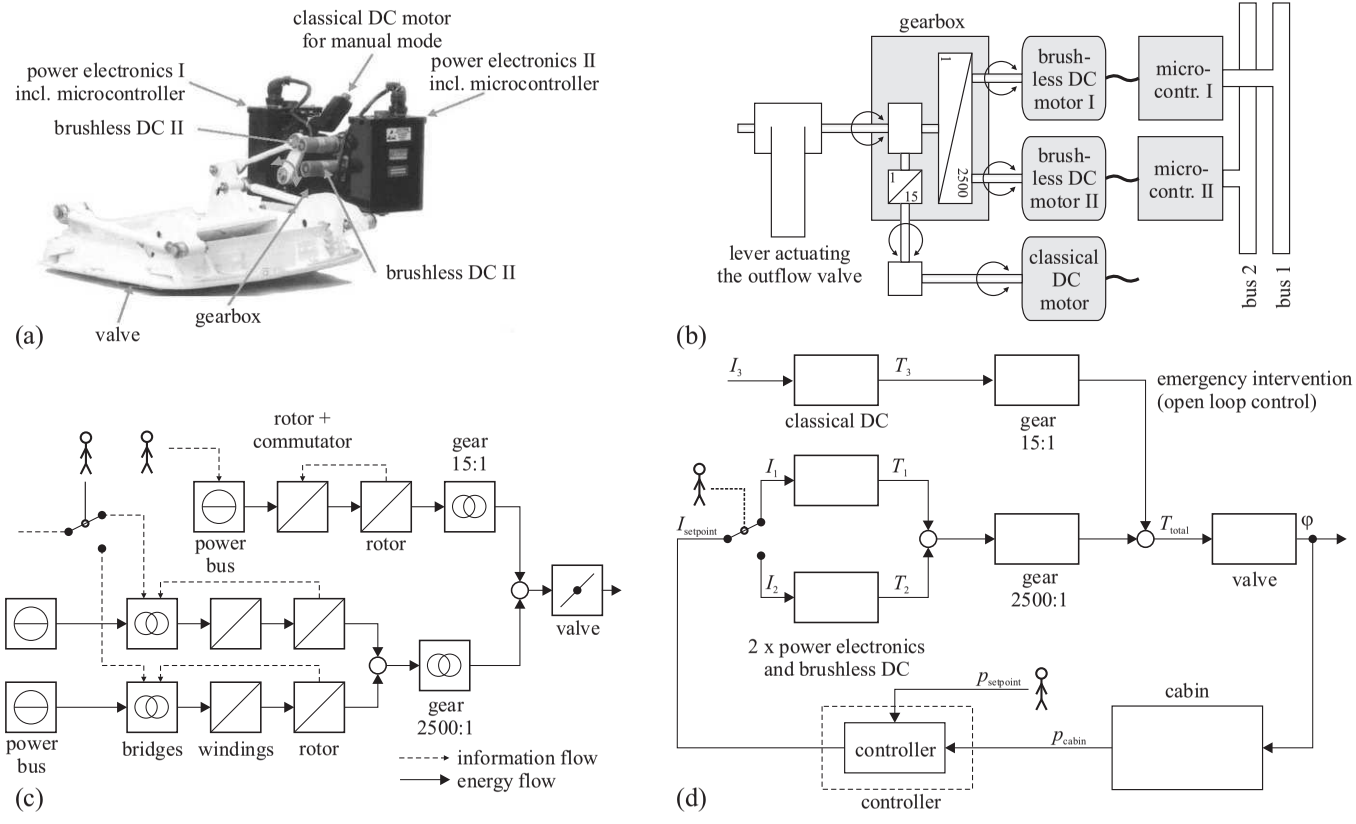


Fig. 12.13. Fault-tolerant electrically actuated cabin outflow valve: a) system; b) schematic; c) energy flow scheme; d) block diagram (redundancy by multiple outflow valves and onboard computers is not shown) [12.32]

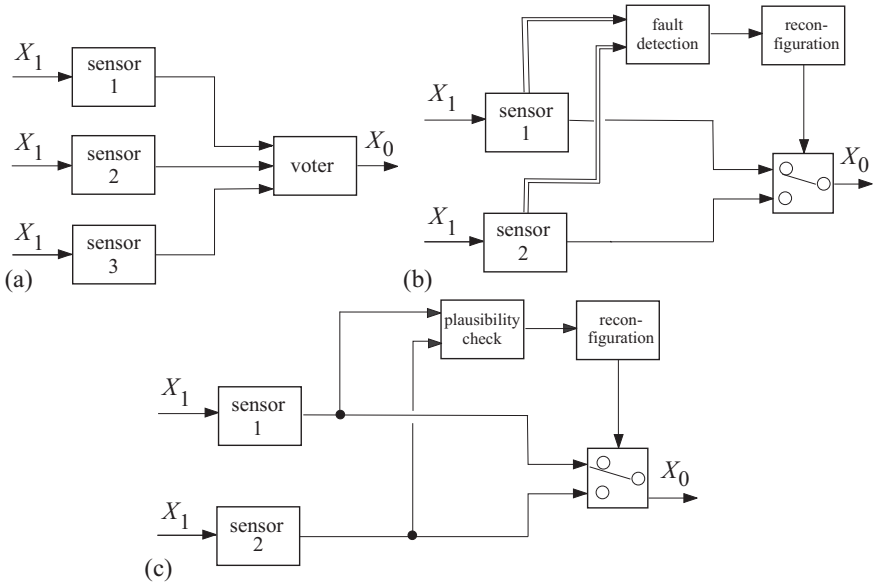


Fig. 12.14. Fault-tolerant sensors with hardware redundancy: a) triplex system with static redundancy and hot standby; b) duplex system with dynamic redundancy, hot standby; c) duplex system with dynamic redundancy, hot standby and plausibility checks

For a process with only one output sensor y_1 and one input sensor u, the output $\hat{y}_1$ can be reconstructed if the process model G_{M1} is known, Figure 12.15b). In both cases, the relationship between the signals of the process are used and expressed in the form of analytical models.

To obtain one usable fault-tolerant measurement value y_{1FT}, at least three different values for y, e.g. the measured one and two reconstructed ones, must be available. This can be obtained by combining the schemes of Figure 12.15a) and b) as shown in Figure 12.16a). A sensor fault y_1 is then detected and masked by a majority voter and either $\hat{y}_1$ or $\hat{y}_{1u}$ is used as a replacement depending on a further decision. (Also, single sensor faults in y_2 or u are tolerated with this scheme.)

One example for this combined analytical redundancy is the yaw rate sensor for the ESP (electronic stability program) of vehicles, where additionally the steering wheel angle as input can be used to reconstruct the yaw rate through a vehicle model as in Figure 12.15b), and the lateral acceleration and the wheel speed difference of the right and left wheel (no slip) are used to reconstruct the yaw rate according to Figure 12.16a).

A more general sensor fault-tolerant system can be designed if two output sensors and one input sensor yield measurements of the same quality. Then, three residuals can be generated and by a decision logic, fault-tolerant outputs can be obtained in the case of single faults of any of the three sensors. This is described in [12.19], Chapter 19. The residuals are generated based on parity equations. In this case, state observers can also be used for residual generation, compare, e.g. the dedicated ob-

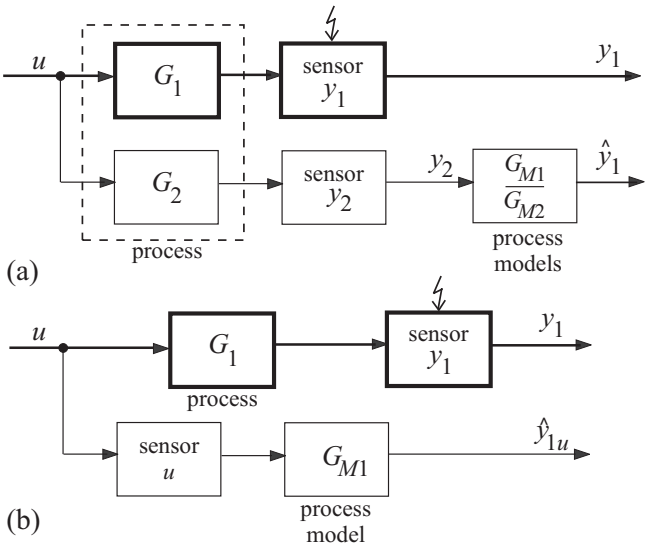


Fig. 12.15. Sensor fault tolerance for one output signal y_1 (main sensor) through analytical redundancy by process models (basic schemes): a) two measured outputs, no measured input; b) one measured input and one measured output. G_i : $G_i(s)$ transfer functions.

servers by [12.7]. (Note that all schemes assume ideal cases. For the realizibility, constraints and additional filters have to be considered.)

If possible, a faulty sensor should be fail-silent, i.e. should be switched off. However, this needs additional switches that lower the reliability. For both hardware and analytical sensor redundancy without fault detection for individual sensors, at least three measurements must be available to make one sensor fail-operational. However, if the sensor (system) has in-built fault detection (integrated self-test or self-validating), two measurements are enough and a scheme like Figure 12.14b) can be applied. (This means that by methods of fault detection, one element can be saved).

Examples of fault-tolerant sensor systems are described in the following.

12.4.3 Steering angle sensor

A first considered realization of a fault-tolerant sensor is the steering angle sensor, see Figure 12.17, [12.12], [12.46]. Information about the steering angle is required for many driver assistance systems, such as the electronic stability control and lane departure warning. Therefore, the reliable measurement of the steering angle is important for the correct action of vehicle control systems.

The basis of the sensor are two GMR (giant magneto resistance) measuring bridges. The axle of the steering wheel is equipped with a gear wheel which drives two pinions. Each pinion turns a permanent magnet whose position is then sensed by one of the two GMR sensors. Each sensor can measure a displacement in the interval $-90°to90°$ with a resolution of 0.1°. The two sensors use the Nonius/Vernier

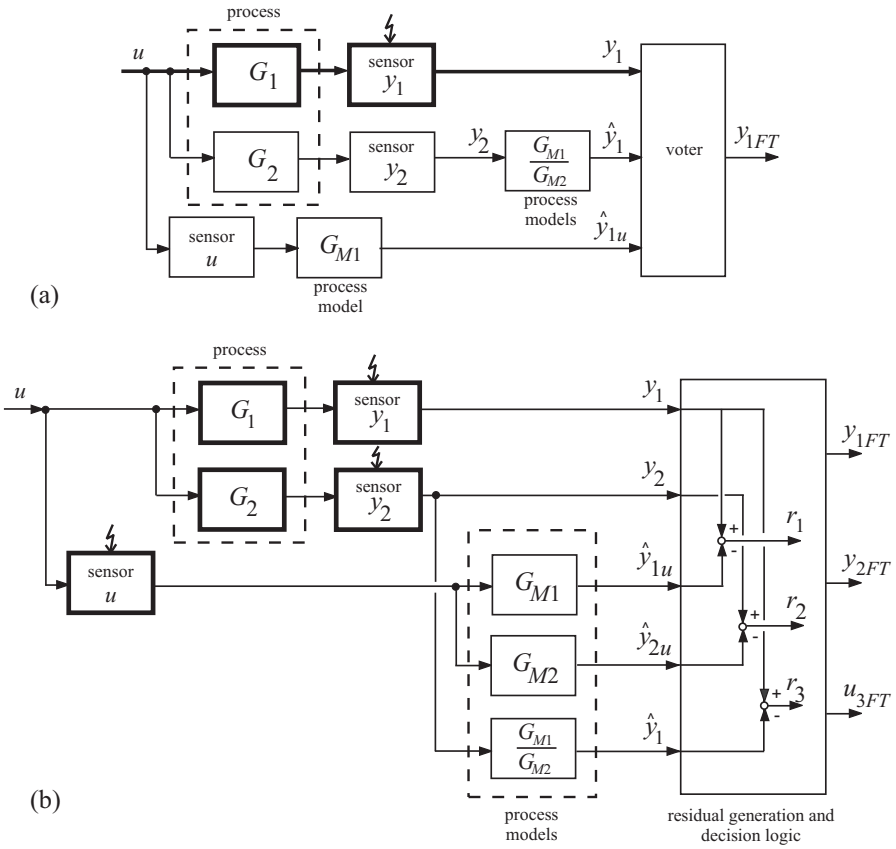


Fig. 12.16. Fault-tolerant sensors with combined analytical redundancy for two measured outputs and one measured input through (analytical) process models: a) y_1 is the main measurement, y_2, u are auxiliary measurements (combination of Figure 12.15a) and b); b) y_1, y_2 and u are measurements of same quality (parity equation approach)

principle to be able to measure the absolute position precisely up to $\pm 720°$. Once the absolute position has been established (e.g. after the start of the engine), the information provided by only one of the two sensors is sufficient to determine relative changes in the position. The evaluation of the sensors along with fault detection is integrated into two separate micro-controllers, which supervise each other. During normal operation, the master is connected to the bus and the slave is only employed the monitor the master. Upon a fault, the master switches off and the slave is connected to the bus, which is symbolized by the switch. The system is FO-FSIL as it can sustain one fault at a sensor or micro-controller without major impairments, only the accuracy of the position reading becomes a bit less accurate. Upon a loss of the second GMR, the sensor module cannot work any longer and switches itself silent.

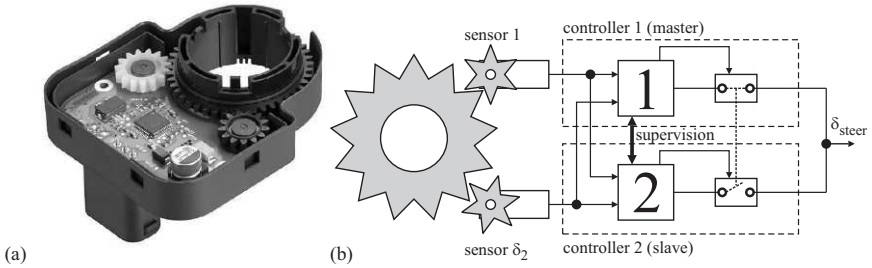


Fig. 12.17. Fault-tolerant steering angle sensor: a) system; b) schematic and block diagram, [12.46]

12.4.4 Fault-tolerant flow sensor

Figure 12.18 presents a fault-tolerant flow sensor, which employs two diverse measuring principles for the measurement of the flow. First, the flow velocity of the fluid is determined by a *vortex generator*. The vortex generator is used to generate eddies in its wash. The vortex shedding leads to vertices, which in turn cause pressure pulsations with a frequency that is proportional to the volume flow rate. The pulsations are either sensed by a pressure sensor or by a force transducer that is connected to the vortex generator. The second measuring principle is based on a differential pressure measurement: The pressure drop between the unconstricted and the constricted part of the pipe is a measure for the fluid velocity, according to Bernoulli's law. The idea behind the fault-tolerant flow rate sensor is that the differential pressure sensor can also be used to detect the higher-frequency pressure pulsations induced by the vertices in the wake of the damming body as they only act on one of the two pressure-sensing lines.

The signal provided by the differential pressure sensor is therefore subject to low-pass filtering in order to determine the average differential pressure between the unconstricted point in the wake of the vortex generator and the point in the constricted area with the Ventouris jet. The signal is at the same time supplied to a high-pass filter which allows one to separate the high-frequency pressure oscillations that are induced by the vertices. A subsequent Fourier transform allows one to determine the frequency of the vortex formation. An integrated fault-detection algorithm allows to detect faults and attribute them to one of the two measuring principles, thereby allowing a consolidation of the measurements, [12.39]. The two diverse measuring principles have different accuracies in different areas of the measuring interval. Thus, upon the loss of one sensing principle, the sensor accuracy may drop. The most critical part is the differential pressure sensor. If this sensor fails, the entire sensor module will be out of operation. Therefore, the system is only FO-FSIL with respect to the sensing principles and, e.g. the metering channels.

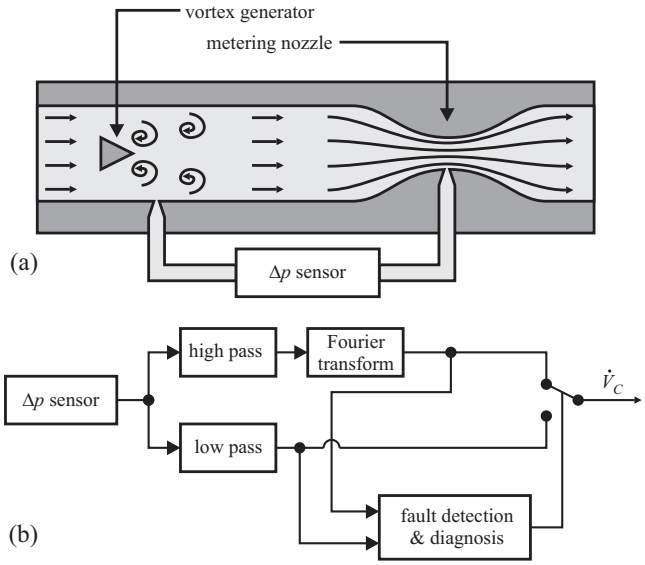


Fig. 12.18. Fault-tolerant flow rate sensor: a) system; b) schematic and block diagram, [12.39]

12.4.5 Electronic throttle

Another example for a fault-tolerant design is the fault-tolerant electronic throttle, see Figure 12.19. The electric motor, a classical DC motor with brushes, may be monitored using model-based methods. The physical model of the motor along with the mechanics also allows one to monitor the two potentiometers and enables the system to decide which of the two potentiometer readings may be faulty, see Section 4.2. Using dynamic redundancy concepts, the system is thus capable of withstanding one sensor fault despite the use of only two instead of three sensors, as would have been necessary in the case of static redundancy, see [12.45]. Figure 12.20 depicts the redundancy concept and Figure 12.21 shows a reconfiguration of a stuck fault of a potentiometer in closed loop, [12.18], [12.45]. The system is active fail-safe (FS) with respect to motor faults. Return springs center the throttle valve in a slightly open position upon a loss of the motor torque (limp-home position). With respect to sensor faults, the system is FO-FS. It can sustain one sensor fault. After the second sensor fault, the system is switched off and the return springs once again bring the throttle into the slightly open position.

12.4.6 Virtual drive dynamic sensors by model based analytical redundancy

Figure 12.22 shows a fault-tolerance approach at the system level, not the component level. Here, the signals from different sensors are brought together, with the aim of providing consolidated measurements to the drive dynamics controllers. The overall design can also be applied to the development of fault-tolerant information platforms for other applications.

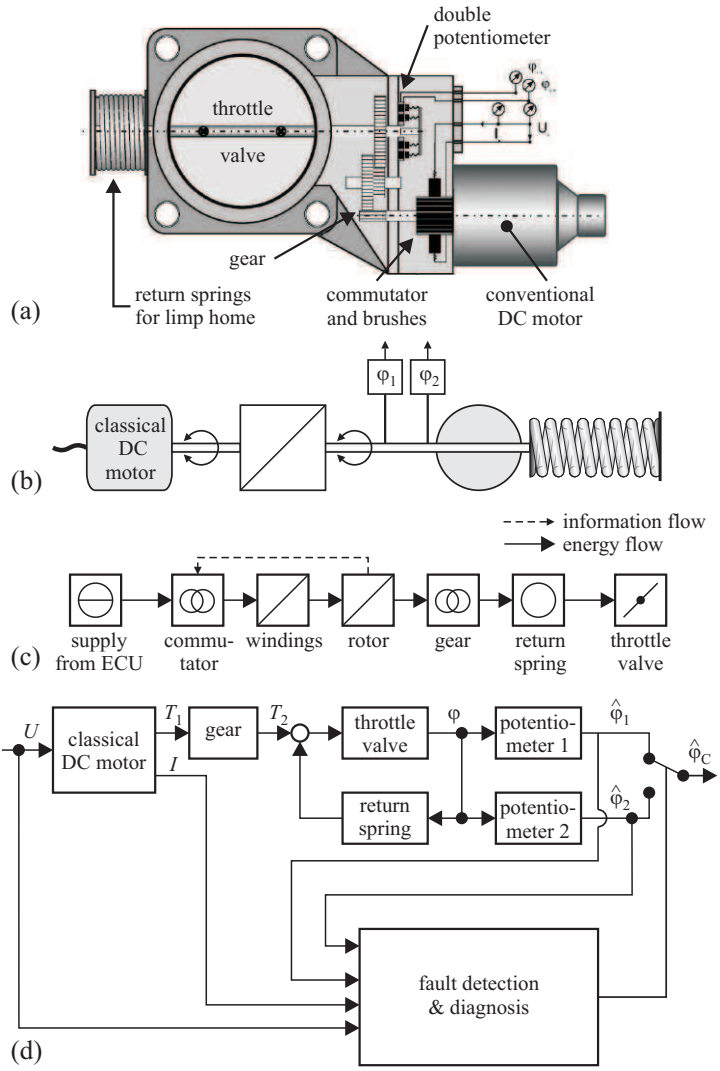


Fig. 12.19. Fault-tolerant electronic throttle: a) throttle plate with gear and DC motor; b) schematic; c) energy flow scheme; d) signal flow scheme [12.45]

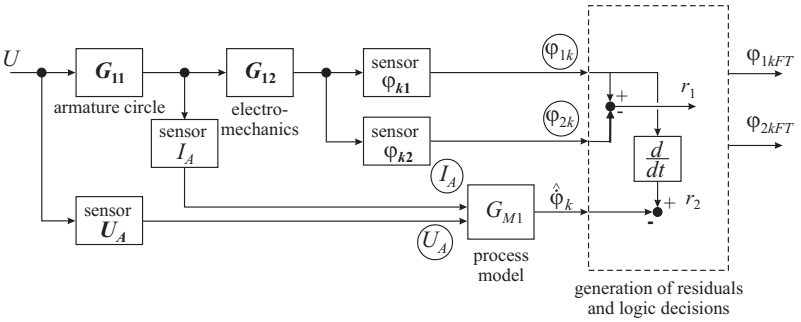


Fig. 12.20. Fault-tolerant position sensor system with two potentiometers and a virtual, model-based position value

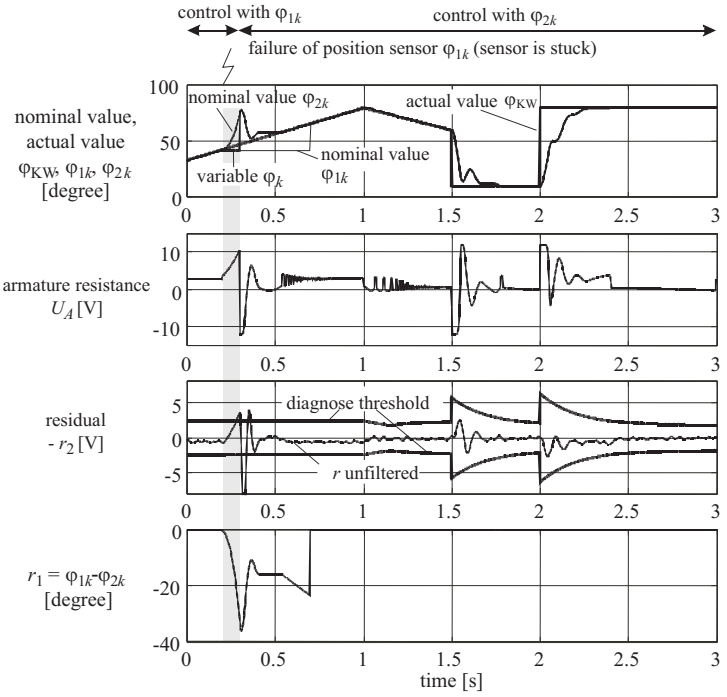


Fig. 12.21. Position control with stuck sensor $1k$ and reconfiguration with the second sensor $2k$ in closed loop

First, the signals are preprocessed, then quantities which cannot be measured or which can only be sensed with expensive or complicated equipment are reconstructed. The reconstruction is based on drive dynamic models of different complexity and employs observes, Kalman filters, parameter estimation algorithms, and local linear neural nets. The fault detection and diagnosis is another important block as all sensors are supervised to identify faulty sensors, which are then masked and replaced by estimated quantities in the block on fault-tolerance methods. Finally, the information platform provides the extracted information to the drive dynamics controllers. These controllers can now not only control directly measurable quantities, but can also control other, reconstructed and more informative quantities, see [12.16]. The drive-dynamics information platform is is *at least* FO-FSIL, but in many cases can sustain more than one sensor fault.

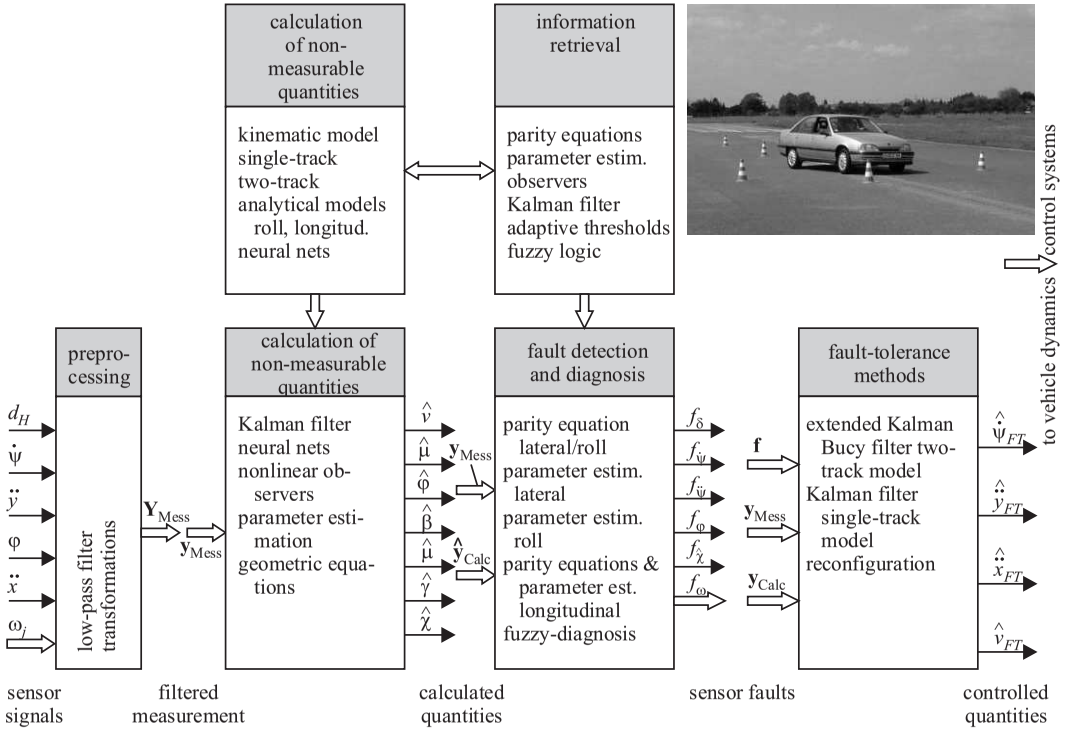


Fig. 12.22. Virtual drive dynamics sensor system, [12.16]. A fault-tolerant sensor platform with analytical redundancy for yaw rate $\dot{\psi}$, lateral acceleration $\ddot{y}$, longitudinal acceleration $\ddot{x}$, over ground velocity v

Part V

Appendix

13

Terminology in fault detection and diagnosis

The following definitions are the result of a coordinated action within the IFAC Technical Committee SAFEPROCESS, published in [13.3]. Some basic definitions can also be found in [13.1], [13.4] and in German standards like DIN and VDI/VDE-Richtlinien, see references at the end of this section and [13.2].

(i) **States and signals**

Fault:	*Unpermitted deviation* of at least one characteristic property of the system
Failure:	*Permanent* interruption of a systems ability to perform a required function under specified operating conditions
Malfunction:	*Intermittent irregularity* in fulfilment of a systems desired function
Error:	*Deviation* between a computed value (of an output variable) and the true, specified or theoretically correct value
Disturbance:	An *unknown* (and uncontrolled) *input* acting on a system
Perturbation:	An input acting on a system which results in a *temporary departure* from a steady state
Residual:	*Fault indicator*, based on deviations between measurements and model-equation-based calculations
Symptom:	*Change* of an observable quantity from *normal behavior*.

(ii) **Functions**

Fault detection:	Determination of faults present in a system and time of detection
Fault isolation:	Determination of kind, location and time of detection of a fault by evaluating symptoms. Follows fault detection
Fault identification:	Determination of the size and time-variant behavior of a fault. Follows fault isolation

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0_13, © Springer-Verlag Berlin Heidelberg 2011

Fault diagnosis:	Determination of kind, size, location and time of detection of a fault by evaluating symptoms. Follows fault detection. Includes fault detection, isolation and identification
Monitoring:	A continuous real-time task of determining the possible conditions of a physical system, recognizing and indicating anomalies of the behavior
Supervision:	Monitoring a physical system and taking appropriate actions to maintain the operation in the case of faults
Protection:	Means by which a potentially dangerous behavior of the system is suppressed if possible, or means by which the consequences of a dangerous behavior are avoided.

(iii) **Models**

Quantitative model:	Use of static and dynamic relations among system variables and parameters in order to describe a system's behavior in quantitative mathematical terms
Qualitative model:	Use of static and dynamic relations among system variables and parameters in order to describe system's behavior in qualitative terms such as causalities or if-then rules
Diagnostic model:	A set of static or dynamic relations which link specific input variables – the symptoms – to specific output variables – the faults
Analytical redundancy:	Use of two, not necessarily identical ways to determine a quantity where one way uses a mathematical process model in analytical form.

(iv) **System properties**

Reliability	Ability of a system to perform a required function under stated conditions, within a given scope, during a given period of time. Measure: MTTF = mean time to failure. MTTF = $1/\lambda$ λ is rate of failure (e.g. failures per hour)
Safety:	Ability of a system not to cause danger to persons or equipment or the environment
Availability:	Probability that a system or equipment will operate satisfactorily and effectively at any point in time. Measure:

$$A = \frac{\text{MTTF}}{\text{MTTF+MTTR}}$$

MTTR mean time to repair
MTTR $= 1/\mu; \mu$: rate of repair

References on terminology

DIN 25424 *Fehlerbaumanalyse (fault tree analysis)*. Beuth Verlag, Berlin, 1990.

DIN 31051 *Instandhaltung (Maintenance)*. Beuth Verlag, Berlin, 1985.

DIN 40041 *Zuverlässigkeit in der Elektrotechnik (Reliability in electrical engineering)*. Beuth Verlag, Berlin, 1990.

DIN 40042 *Zuverlässigkeit elektrischer Geräte, Anlagen und Systeme (Reliability of electrical devices, plants and systems)*. Beuth Verlag, Berlin, 1989.

DIN 55350 *Begriffe der Qualitätssicherung und Statistik (Terms in quality control and statistics)*. Beuth Verlag, Berlin, 1989.

IFIP working group 10.4. Reliable computing and fault tolerance, meeting in Como, Italy, 1983.

Laprie, J.C. (1983). *On computer system dependability and un-dependability: faults, errors, and failures*. IFIP WG 10.4, Como, Italy, 1983.

Lexikon Mess- und Automatisierungstechnik. (1992). VDI Verlag, Düsseldorf.

Reliability, Availability, and Maintainability Dictionary. ASQC Quality Press, Milwaukee, WI, 1988.

Robinson, A. (1982). A user-oriented perspective of fault-tolerant systems, models and terminologies. *Proceedings of the 12th International Symposium on Fault-Tolerant Computing*, Los Angeles.

VDI/VDE-Richtlinie 3541. *Steuerungseinrichtungen mit vereinbarter gesicherter Funktion*. Beuth Verlag, Berlin, 1985.

VDI/VDE-Richtlinie 3542. *Sicherheitstechnische Begriffe für Automatisierungssysteme*. Beuth Verlag, Berlin, 1988.

VDI/VDE-Richtlinie 3691. *Erfassung von Zuverlässigkeitswerten bei Prozessrechnereinsätzen*. Beuth Verlag, Berlin, 1985.

Concluding remarks

Based on the fault-detection and fault-diagnosis methods treated in the book **Fault-Diagnosis Systems** (2006), Springer-Verlag, this book has shown how the different methods can be applied to the various technical processes. The selection of different technical processes shows by the use of *process-model-based* and *signal-model-based* methods how several analytical symptoms can be generated to detect a larger number of faults, especially in an earlier fault-development phase and to diagnose the faults.

Process-model-based fault detection can be realized with the aid of *process models* and several measurements. The application of process-model-based fault-detection methods requires that at least one input and one corresponding output signal can be measured. The additional measurement of signals in the signal flow between the input and output signals may support the fault-detection capability. The process models used have to be relatively precise. Herewith it is recommended that the model structure is obtained from a physical/theoretical modeling approach and that the parameters of the used process model are estimated by experiments with parameter estimation methods. Linear dynamic models can be used quite often, especially for stationary operating states. The kind of nonlinear models results mostly from theoretical modeling based on balance equations, constitutive equations and phenomenological laws.

For processes in *stationary operation* around a fixed operating point especially parity equations can be successfully applied. Examples are components of industrial plants like pumps, heat exchangers, and pipelines. In the case of *nonstationary, dynamic operation* parameter-estimation methods allow the detection and the differentiation of a larger number of faults, also if only a few measurable signals are available. This holds, for example for drives, actuators, robots, pumps, machine tools, heat exchangers and vehicles. In combination with parity equations several symptoms can be generated which enable a detailed fault diagnosis. Parameter estimation requires in general an appropriate excitation of an input signal, either from normal operation or artificially introduced. Parity equations or state observers can also be applied if the input signals do not change.

R. Isermann, *Fault-Diagnosis Applications, Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems,*
DOI 10.1007/978-3-642-12767-0, © Springer-Verlag Berlin Heidelberg 2011

Fault-detection methods based on *signal models* may be based on periodical or stochastic measured signals. It is then possible to analyze single sensor outputs. Processes with periodic operation cycles are for example oscillating reciprocating pumps, combustion engines and some machine tools. Then, periodical signal models can be applied. Changes induced through faults can be detected by the estimation of amplitudes, phases and frequencies with the methods of Fourier or wavelet analysis or by band-pass filters. If only a few unknown frequencies have to be estimated an ARMA signal parameter estimation can be used. A combination of signal-model analysis with process-model-based methods, e.g. parity equations for mean value process models increases the number of symptoms and therefore the fault coverage like for combustion engines or for reciprocating pumps.

Dynamic state-space observers could be successfully applied for the leak detection of gas pipelines as the leaks appear as state-variable changes of a pipeline section. However, if state variables are not changed directly by faults, e.g. by parameter changes, then state-space observers are usually not suitable. Fault-sensitive observers, dedicated observers or output observers and corresponding Kalman filters which are specially designed for fault detection assume generally multi-variable process models. They often yield about the same results as the simpler parity equations. Linear and nonlinear state space observers are, however, very well suited to determine non-measurable values as the slip angle for vehicles or stresses and temperatures in mechanical components.

A special problem is the fault detection in and for *closed loops* as they compensate small faults of actuators, sensors and processes. Therefore, they are difficult to detect. Only larger faults change the control performance significantly. However, similar changes of the control behavior arise for larger disturbances or for not well-tuned controllers. Then a combination of several detection methods is required.

The table on the next page gives a *summary of the application of model-based fault-detection methods* for the processes treated in this book. It depicts the kind of operation, the measured variables, applied linear or nonlinear process models, the kind of identification, the use of the different fault-detection methods, and the number of detected faults.

A *fault diagnosis* requires in general the generation of several significant symptoms. In this book mostly fault-symptom tables have been used as a common easy to interpret representation. A differentiation (isolation) of special faults is then possible if the signs and the values of the symptoms are significantly different. Then in many cases a simple pattern recognition is sufficient to be used for classification. For a systematic treatment of fault-symptom trees with a following up of different decisions the approximative reasoning with if-then-rules by fuzzy logic is a feasible way and was especially successful. The application of this inference-based method was described for some processes.

In the case of *electrical driven drive trains* directly measurable variables of electrical motors like voltage, current and speed can be used for the fault detection of the electrical motor and additionally also for the connected mechanical parts like gears and any machinery. This can be called “drive-as-sensor principle” and was demonstrated, e.g. for electrical actuators, pumps and machine tools.

Application of model-based fault-detection methods (IAT, TU Darmstadt) X successful, (x) not recommended, [3*p*: three pressure sensors]

Process	Operation mode				Measure-ments	Process models									Signal models		No. of detected faults
	sta-tio-nary	non-statio-nary	open loop	closed loop		linear	non-linear	identification			parity equations		observer		non-parametr.	para-metr.	
								corre-lation	param. estim.	net models	output error	equat. error	state observer	fault observer		ARMA	
Electromotor																	
– DC	X	X	X		U, I, ω	X	X		X			X	(x)	(x)			7
– AC		X		X	U, I, ω		X		X	X	X						> 9
Actuator																	
– electrical		X	X		U, I, ω, φ	X	X		X		X						5
– pneumatic		X		X	$w, z, \dot{v}$						X						6
– hydraulic		X		X	$u, z, 3p$		X		X	(x)	X						16
Industrial robot		X		X	U, I, ω, φ		X		X								6
Machine tools																	
– feed drive		X		X	U, I, ω, z		X		X						Fourier-Tr.		7
– milling				X	φ_1, φ_2, z		X		X						transf.		3
– grinding	X	X		X	$I, \dot{z}$	X			X						(x)	X	3
Pipeline																	
– liquid	X	X	X		$2\dot{m}, 2p$	X		X									1
– gas	X	X	X		$2\dot{m}, 2p$		X						X				1
Pumps																	
– centrif. (DC)	X	X	X		$U, I, \Delta p, \dot{V}$		X		X								15
– centrif. (AC)			X		$I, \omega, \Delta p, \dot{V}$					X	X				Bandpass		9
– diaphragm	X		X		p, z, φ		X				X						4
Heat exchanger																	
– steam/water	X		X		$2\dot{m}, 2T$	X			X								4
– steam/water		X	X		$2\dot{m}, 2T$		X			X	X						9

The last chapters of this book have shown that another way to cope with appearing faults is the design of *fault-tolerant systems*, for example for drives, actuators and sensors or certain process components. This requires generally a redundancy and a possibility to reconfigure the system after the fault detection of a component.

More results for the fault diagnosis of *internal combustion engines* and *automobiles* will be published in separate books.

References

Chapter 1

1.1 Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. *Diagnosis and fault-tolerant control*. Springer, Berlin, 2nd edition, 2006.

1.2 Chen, J. and Patton, R. *Robust model-based fault diagnosis for dynamic systems*. Kluwer, Boston, 1999.

1.3 Gertler, J. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, New York, 1998.

1.4 Himmelblau, D. *Fault detection and diagnosis in chemical and petrochemical processes*. Elsevier, New York, 1978.

1.5 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.

1.6 NAMUR-recommendation NE 107. *Self-monitoring and diagnosis of field devices. www.NAMUR.de*. NAMUR, Leverkusen, 2005.

1.7 NAMUR-recommendation NE 91. *Requirements for online asset management. www.NAMUR.de*. NAMUR, Leverkusen, 2001.

1.8 Patton, R. Fault-tolerant control: the 1997 situation. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 2, pages 1033–1055, Hull, UK, August 1997. Pergamon Press.

1.9 Patton, R., Frank, P., and Clark, P., editors. *Fault diagnosis in dynamic systems, theory and application*. Prentice Hall, London, 1989.

Chapter 2

2.1 Barlow, R. and Proschan, F. *Statistical theory of reliability and life testing*. Holt, Rinehart & Winston, 1975.

2.2 Beard, R. Failure accommodation in linear systems through self-reorganization. Technical Report MVT-71-1, Man Vehicle Laboratory, Cambridge, MA, 1971.

2.3 Bonnett, A. Understanding motor shaft failures. *IEEE Industry Application Magazine*, (September–October):25–41, 1999.
2.4 Chen, J. and Patton, R. *Robust model-based fault diagnosis for dynamic systems*. Kluwer, Boston, 1999.
2.5 Clark, R. State estimation schemes for instrument fault detection. In Patton, R., Frank, P., and Clark, R., editors, *Fault diagnosis in dynamic systems*, chapter 2, pages 21–45. Prentice Hall, New York, 1989.
2.6 Clark, R. A simplified instrument detection scheme. *IEEE Trans. Aerospace Electron. Systems*, 14(3):558–563, 1990.
2.7 Dalton, T., Patton, R., and Chen, J. An application of eigenstructure assignement to robust residual design for FDI. In *Proc. UKACC Int. Conf. on Control (CONTROL´96)*, pages 78–83, Exeter, UK, 1996.
2.8 Ericsson, S., Grip, N., Johannson, E., Persson, L., Sjöberg, R., and Strömberg, J. Towards automatic detection of local bearing defects in rotating machines. *Mechanical Systems and Signal Processing*, 9:509–535, 2005.
2.9 Filbert, D. Fault diagnosis in nonlinear electromechanical systems by continuous-time parameter estimation. *ISA Trans.*, 24(3):23–27, 1985.
2.10 Frank, P. Advanced fault detection and isolation schemes using nonlinear and robust observers. In *10th IFAC Congress*, volume 3, pages 63–68, München, Germany, 1987.
2.11 Frank, P. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy. *Automatica*, 26(3):459–474, 1990.
2.12 Frank, P. Enhancement of robustness in observer-based fault detection. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 1, pages 275–287, Baden-Baden, Germany, September 1991. Pergamon Press.
2.13 Freyermuth, B. *Wissensbasierte Fehlerdiagnose am Beispiel eines Industrieroboters*. Fortschr.-Ber. VDI Reihe 8, 315. VDI Verlag, Düsseldorf, 1993.
2.14 Frost, R. *Introduction to knowledge base systems*. Collins, London, 1986.
2.15 Füssel, D. *Fault diagnosis with tree-structured neuro-fuzzy systems*. Fortschr.-Ber. VDI Reihe 8, 957. VDI Verlag, Düsseldorf, 2002.
2.16 Füssel, D. and Isermann, R. Hierarchical motor diagnosis utilizing structural knowledge and a self-learning neuro-fuzzy-scheme. *IEEE Trans. on Ind. Electronics*, 74(5):1070–1077, 2000.
2.17 Geiger, G. *Technische Fehlerdiagnose mittels Parameterschätzung und Fehlerklassifikation am Beispiel einer elektrisch angetriebenen Kreiselpumpe*. Fortschr.-Ber. VDI Reihe 8, 91. VDI Verlag, Düsseldorf, 1985.
2.18 Gertler, J. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, New York, 1998.
2.19 Grimmelius, H., Meiler, P., Maas, H., Bonnier, B., Grevink, J., and Kuilenburg, R. van. Three state-of-the-art methods for condition monitoring. *IEEE Trans. on Industrial Electronics*, 46(2):401–416, 1999.
2.20 Hermann, O. and Milek, J. Modellbasierte Prozessüberwachung am Beispiel eines Gasverdicherts. *Technisches Messen*, 66(7–8):293–300, 1995.

2.21 Higham, E. and Perovic, S. Predictive maintenance of pumps based on signal analysis of pressure and differential pressure (flow) measurements. *Trans. of the Institute of Measurement and Control*, 23(4):226–248, 2001.
2.22 Himmelblau, D. *Fault detection and diagnosis in chemical and petrochemical processes*. Elsevier, New York, 1978.
2.23 Himmelblau, D. Fault detection and diagnosis - today and tomorrow. In *Proc. IFAC Workshop on Fault Detection and Safety in Chemical Plants*, pages 95–105, Kyoto, Japan, 1986.
2.24 Höfling, T. *Methoden zur Fehlererkennung mit Parameterschätzung und Paritätsgleichungen*. Fortschr.-Ber. VDI Reihe 8, 546. VDI Verlag, Düsseldorf, 1996.
2.25 Höfling, T. and Isermann, R. Fault detection based on adaptive parity equations and single-parameter tracking. *Control Engineering Practice – CEP*, 4(10):1361–1369, 1996.
2.26 IEC 61508. *Functional safety of electrical/electronic/programmable electronic systems*. International Electrotechnical Commission, Switzerland, 1997.
2.27 IFIP. *Proc. of the IFIP 9th World Computer Congress, Paris, France, September 19–23*. Elsevier, 1983.
2.28 Isermann, R. Process fault detection on modeling and estimation methods - a survey. *Automatica*, 20(4):387–404, 1984.
2.29 Isermann, R. Estimation of physical parameters for dynamic processes with application to an industrial robot. *International Journal of Control*, 55(6):1287–1298, 1992.
2.30 Isermann, R. Integration of fault-detection and diagnosis methods. In *Proc. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, pages 597–609, Espoo, Finland, June 1994.
2.31 Isermann, R., editor. *Überwachung und Fehlerdiagnose - Moderne Methoden und ihre Anwendungen bei technischen Systemen*. VDI-Verlag, Düsseldorf, 1994.
2.32 Isermann, R. Supervision, fault-detection and fault-diagnosis methods – an introduction. *Control Engineering Practice – CEP*, 5(5):639–652, 1997.
2.33 Isermann, R. Diagnosis methods for electronic controlled vehicles. *Vehicle System Dynamics*, 36(2-3):77–117, 2001.
2.34 Isermann, R. Fehlertolerante Komponenten für Drive-by-wire Systeme. *Automobiltechnische Zeitschrift – ATZ*, 104(4):382–392, 2002.
2.35 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.
2.36 Isermann, R. Model-based fault detection and diagnosis – status and applications. *Annual Reviews in Control*, 29:71–85, 2005.
2.37 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
2.38 Isermann, R. and Ballé, P. Trends in the application of model-based fault detection and diagnosis in technical processes. *Control Engineering Practice – CEP*, 5(5):638–652, 1997.

2.39 Isermann, R. and Freyermuth, B. Process fault diagnosis based on process model knowledge. *Journal of Dynamic Systems, Measurement and Control*, 113:Part I, 620–626; Part II, 627–633, 1991.
2.40 Isermann, R., Lachmann, K.-H., and Matko, D. *Adaptive control systems*. Prentice Hall International UK, London, 1992.
2.41 ISO 13374. *Condition monitoring and diagnostics of machines – date processing, communciation, and presenation. Draft*. International Organization for Standardization, Geneva, 2005.
2.42 ISO 13374-2. *Condition monitoring and diagnostics of machines – Data processing, communication and presentation – Part 2: Data processing*. International Organization for Standardization, Geneva, 2007.
2.43 Jones, H., editor. *Failure detection in linear systems*. Dept. of Aeronautics, M.I.T., Cambridge, 1973.
2.44 Kiencke, U. Diagnosis of automotive systems. In *Proc. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, Hull, UK, August 1997. Pergamon Press.
2.45 Kolerus, J. *Zustandsüberwachung von Maschinen*. expert Verlag, Renningen-Malmsheim, 2000.
2.46 Leonhardt, S. *Modellgestützte Fehlererkennung mit neuronalen Netzen - Überwachung von Radaufhängungen und Diesel-Einspritzanlagen*. Fortschr.-Ber. VDI Reihe 12, 295. VDI Verlag, Düsseldorf, 1996.
2.47 Melody, J., Basar, T., Perkins, W., and Voulgaris, P. Parameter estimation for inflight detection of aircraft icing. In *Proc. 14th IFAC World Congress*, pages 295–300, Beijing, P.R. China, 1991.
2.48 Musgrave, J., Guo, T.-H., Wong, E., and Duyar, A. Real-time accommodation of actuator faults on a reusable rocket engine. *IEEE Trans. on Control Systems Technology*, 5(1):100–109, 1997.
2.49 Nold, S. *Wissensbasierte Fehlererkennung und Diagnose mit den Fallbeispielen Kreiselpumpe und Drehstrommotor*. Fortschr.-Ber. VDI Reihe 8, 273. VDI Verlag, Düsseldorf, 1991.
2.50 Omdahl, T., editor. *Reliability, availability and maintainability (RAM) dictionary*. ASQC Quality Press, Milwaukee, WI, USA, 1988.
2.51 Patton, R. Fault detection and diagnosis in aerospace systems using analytical redundancy. *IEE Computing & Control Eng. J.*, 2(3):127–136, 1991.
2.52 Patton, R., Frank, P., and Clark, P., editors. *Issues of fault diagnosis for dynamic systems*. Springer, New York, 2000.
2.53 Rasmussen, J. Diagnostic reasoning in action. *IEEE Trans. on System, Man and Cybernetics*, 23(4):981–991, 1993.
2.54 Rizzoni, G., Soliman, A., and Passino, K. A survey of automotive diagnostic equipment and procedures. SAE 930769. In *Proc. International Congress and Exposition*, Detroit, MI, USA, 1993. SAE.
2.55 Russell, E., Chiang, L., and Baatz, R. *Data-driven techniques for fault detection and diagnosis in chemical processes*. Springer, London, 2000.

2.56 Schneider-Fresenius, W. *Technische Fehlerfrühdiagnose-Einrichtungen: Stand der Technik und neuartige Einsatzmöglichkeiten in der Maschinenbauindustrie*. Oldenbourg, München, 1985.
2.57 Sill, U. and Zörner, W. *Steam turbine generators process control and diagnostics – modern instrumentation for the greatest economy of power plants*. Wiley-VCH, Weinheim, 1996.
2.58 Storey, N. *Safety-critical computer systems*. Addison Wesley Longman Ltd., Essex, 1996.
2.59 Struss, P., Malik, A., and Sachenbacher, M. Qualitative modeling is the key to automated diagnosis. In *13th IFAC World Congress*, San Francisco, CA, USA, 1996.
2.60 Sturm, A. and Förster, B. *Maschinen und Anlagendiagnostik*. B.G. Teubner, Stuttgart, 1986.
2.61 Sturm, A., Förster, B., Hippmann, N., and Kinsky, D. *Wälzlaufdiagnose an Maschinen und Anlagen*. Verlag TÜV Rheinland, Köln, 1986.
2.62 Torasso, P. and Console, L. *Diagnostic problem solving*. North Oxford Academic, Oxford, 1989.
2.63 Tou, J. and Gonzalez, R. *Pattern recognition principles*. Addison-Wesley Publishing, Reading, MA, 1974.
2.64 VDMA Fachgemeinschaft Pumpen. *Betreiberumfrage zur Störungsfrüherkennung bei Pumpen*. VDMA, Frankfurt, 1995.
2.65 Wang, L. and Gao, R., editors. *Condition monitoring and control for intelligent manufacturing*. Springer, London, 2006.
2.66 Willsky, A. A survey of design methods for failure detection systems. *Automatica*, 12:601–611, 1976.
2.67 Wolfram, A. *Komponentenbasierte Fehlerdiagnose industrieller Anlagen am Beispiel frequenzumrichtergespeister Asynchronmaschinen und Kreiselpumpen*. Fortschr.-Ber. VDI Reihe 8, 967. VDI Verlag, Düsseldorf, 2002.
2.68 Wowk, V. *Machinery vibrations*. McGraw Hill, New York, 1991.

Chapter 3

3.1 Bünte, A. and Grotstollen, H. Offline parameter identification of an invert-fed induction motor at standstill. In *6 EPE 6th European Conference on Power Electronics and Applications*, Seville, Spain, 1995.
3.2 Filbert, D. Technical diagnosis for the quality control of electrical low power motors (in German). *Technisches Messen*, 70(9):417–427, 2003.
3.3 Fraser, C. and Milne, J. *Electro-mechanical engineering – an integrated approach*. IEEE Press, Piscataway, NJ, 1994.
3.4 Füssel, D. *Fault diagnosis with tree-structured neuro-fuzzy systems*. Fortschr.-Ber. VDI Reihe 8, 957. VDI Verlag, Düsseldorf, 2002.

3.5 Höfling, T. Zustandsgrößenschätzung zur Fehlererkennung. In Isermann, R., editor, *Überwachung und Fehlerdiagnose*, pages 89–108. VDI, Düsseldorf, 1994.
3.6 Höfling, T. *Methoden zur Fehlererkennung mit Parameterschätzung und Paritätsgleichungen*. Fortschr.-Ber. VDI Reihe 8, 546. VDI Verlag, Düsseldorf, 1996.
3.7 Höfling, T. and Isermann, R. Fault detection based on adaptive parity equations and single-parameter tracking. *Control Engineering Practice – CEP*, 4(10):1361–1369, 1996.
3.8 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2003.
3.9 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.
3.10 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
3.11 Isermann, R., Lachmann, K.-H., and Matko, D. *Adaptive control systems*. Prentice Hall International UK, London, 1992.
3.12 Kastha, D. and Bose, B. Investigation of fault modes of voltage-fed inverter system for induction motor. *IEEE Trans. on Industry Applications*, 30(4):1028–1037, 1994.
3.13 Leonhard, W. *Control of electrical drives*. Springer, Berlin, 2nd edition, 1996.
3.14 Lyshevski, S. *Electromechanical systems, electric machines, and applied mechatronics*. CRC Press, Boca Raton, FL, 2000.
3.15 Nelles, O. *Nonlinear system identification*. Springer, Heidelberg, 2001.
3.16 Pfeufer, T. Improvement of flexibility and reliability of automobiles actuators by model-based algorithms. In *IFAC SICICA*, Budapest, Hungary, 1994.
3.17 Pfeufer, T. *Modellgestützte Fehlererkennung und Diagnose am Beispiel eines Fahrzeugaktors*. Fortschr.-Ber. VDI Reihe 8, 749. VDI Verlag, Düsseldorf, 1999.
3.18 Sarma, M. *Electric machines. Steady-state theory and dynamic performance*. PWS Press, New York, 1996.
3.19 Schröder, D. *Elektrische Antriebe 1*. Springer, Berlin, 1995.
3.20 Stölting, H. Electromagnetic actuators. In Janocha, H., editor, *Actuators*. Springer, Berlin, 2004.
3.21 Thorsen, O. and Dalva, M. A survey of the reliability with an analysis of faults on variable frequency drives in industry. In *Proc. European Conference on Power Electronics and Applications EPE '95*, pages 1033–1038, 1995.
3.22 Wolfram, A. *Komponentenbasierte Fehlerdiagnose industrieller Anlagen am Beispiel frequenzumrichtergespeister Asynchronmaschinen und Kreiselpumpen*. Fortschr.-Ber. VDI Reihe 8, 967. VDI Verlag, Düsseldorf, 2002.
3.23 Wolfram, A. and Isermann, R. On-line fault detection of inverter-fed induction motors using advanced signal processing techniques. In *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS´2000)*, Budapest, Hungary, June 2000.

3.24 Wolfram, A. and Isermann, R. Fault detection of inverter-fed induction motors using a multimodel approach based on neuro-fuzzy models. In *Proc. European Control Conference*, Porto, Portugal, September 2001.

Chapter 4

4.1 Ayoubi, M. *Nonlinear system identification based on neural networks with locally distributed dynamics and application to technical processes*. Fortschr.-Ber. VDI Reihe 8, 591. VDI Verlag, Düsseldorf, 1996.
4.2 Gertler, J. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, New York, 1998.
4.3 Höfling, T. *Methoden zur Fehlererkennung mit Parameterschätzung und Paritätsgleichungen*. Fortschr.-Ber. VDI Reihe 8, 546. VDI Verlag, Düsseldorf, 1996.
4.4 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.
4.5 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
4.6 Isermann, R. and Keller, H. Intelligente Aktoren. *atp – Automatisierungstechnische Praxis*, 35:593–602, 1993.
4.7 Janocha, H., editor. *Actuators. Basics and Applications*. Springer, Berlin, 2004.
4.8 Kallenbach, E., Eick, R., Quandt, P., Ströhla, T., Feindt, K., and Kallenbach, M. *Elektromagnete: Grundlagen, Berechnung, Entwurf und Anwendung*. Teubner, Stuttgart, 3rd edition, 2008.
4.9 Moseler, O. *Mikrocontrollerbasierte Fehlererkennung für mechatronische Komponenten am Beispiel eines elektromechanischen Stellantriebs*. Fortschr.-Ber. VDI Reihe 8, 980. VDI Verlag, Düsseldorf, 2001.
4.10 Moseler, O., Heller, T., and Isermann, R. Model-based fault detection for an actuator driven by a brushless DC motor. In *14th IFAC World Congress*, volume P, pages 193–198, Beijing, China, 1999.
4.11 Moseler, O. and Isermann, R. Application of model-based fault detection to a brushless DC motor. *IEEE Trans. on Industrial Electronics*, 47(5):1015–1020, 2000.
4.12 Moseler, O. and Müller, M. A smart actuator with model-based FDI implementation on a microcontroller. In *1st IFAC Conference on Mechatronic Systems*, Darmstadt, Germany, September 2000.
4.13 Moseler, O. and Vogt, M. FIT- filtering and identification. In *Proc. 12th IFAC Symposium on System Identification (SYSID)*, Santa Barbara, CA, USA, 2000.
4.14 Pfeufer, T. *Modellgestützte Fehlererkennung und Diagnose am Beispiel eines Fahrzeugaktors*. Fortschr.-Ber. VDI Reihe 8, 749. VDI Verlag, Düsseldorf, 1999.
4.15 Pfeufer, T., Isermann, R., and Rehm, L. Quality assurance of mechanical-electronical automobile actuator using an integrated model-based diagnosis

control (in German). In *Proc. VDI-Conference Elektronik im Kraftfahrzeug*, volume VDI-Bericht Nr. 1287, pages 145–159, September 1996.

4.16 Raab, U. *Modellgestützte digitale Regelung und Überwachung von Kraftfahrzeugaktoren*. Fortschr.-Ber. VDI Reihe 8, 313. VDI Verlag, Düsseldorf, 1993.

4.17 Streib, H.-M. and Bischof, H. Electronic throttle control (ETC): A cost effective system for improved emissions, fuel economy, and driveability. In *SAE International Congress & Exposition*, number 960338, Warrendale, PA, 1996.

Chapter 5

5.1 An, L. and Sepehri, N. Hydraulic actuator circuit fault detection using extended Kalman filter. In *Proc. American Control Conference*, volume 5, pages 4261–4266, Denver, CO, June 2003.

5.2 Backé, W. *Grundlage der Pneumatik*. RWTH Aachen, Aachen, 7nd edition, 1986.

5.3 Ballé, P. *Modellbasierte Fehlererkennung für nichtlineare Prozesse mit linearparameterveränderlichen Modellen*. Fortschr.-Ber. VDI Reihe 8, 960. VDI Verlag, Düsseldorf, 2002.

5.4 Ballé, P. and Füssel, D. Engineering applications of artificial intelligence. *Control Engineering Practice – CEP*, 13:695–704, 2000.

5.5 Choudhury, M. S., Shah, S., Thornhill, N., and Shook, D. S. Automatic detection and quantification of stiction in control valves. *Control Engineering Practice*, 14(12):1395–1412, 2006.

5.6 Deibert, R. *Methoden zur Fehlererkennung an Komponenten im geschlossenen Regelkreis*. Fortschr.-Ber. VDI Reihe 8, 650. VDI Verlag, Düsseldorf, 1997.

5.7 Füssel, D. *Fault diagnosis with tree-structured neuro-fuzzy systems*. Fortschr.-Ber. VDI Reihe 8, 957. VDI Verlag, Düsseldorf, 2002.

5.8 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2003.

5.9 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.

5.10 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.

5.11 Karpenko, M., Sepehri, N., and Scuse, D. Diagnosis of process valve actuator faults using a multilayer neural network. *Control Engineering Practice – CEP*, 11:1289–1299, 2003.

5.12 Kazemi-Moghaddan, A. *Fehlerfrühidentifikation und -diagnose elektrohydraulischer Linearantriebssysteme*. Doctoral thesis. Technische Universität, Fachbereich Maschinenbau, Darmstadt, 1999.

5.13 Keller, H. *Wissensbasierte Inbetriebnahme und adaptive Regelung eines pneumatischen Linearantriebs*. Fortschr.-Ber. VDI Reihe 8, 412. VDI Verlag, Düsseldorf, 1994.

5.14 Khan, H., Abour, S., and Sepehri, N. Nonlinear observer-based fault-detection techniques for electro-hydraulic servo-positioning systems. *Mechatronics*, 15:1037–1059, 2005.
5.15 Kiesbauer, J. Diagnosetool bei Stellgeräten. *Automatisierungstechnische Praxis – atp*, 42(3):38–45, 2000.
5.16 Kiesbauer, J. Neues integriertes Diagnosekonzept bei digitalen Stellungsreglern. *Automatisierungstechnische Praxis – atp*, 46(4):40–48, 2004.
5.17 Kollmann, E. Wirkung wesentlicher Nichtlinearitäten auf die Stabilität und den statischen Fehler von Stellungsreglern. *Automatik*, (11):379–383, 1968.
5.18 Kress, R. *Robuste Fehlerdiagnoseverfahren zur Wartung und Serienabhnahme elektrohydraulische Aktuatoren*. Doctoral thesis. TU Darmstadt, Fachbereich Maschinenbau, Darmstadt, 2002.
5.19 McGhee, J., Henderson, I., and Baird, A. Neural networks applied for the identification and fault diagnosis of process valves and actuators. *Measurement Journal of Int. Measurm. Conf.*, 20(4):267–275, 1997.
5.20 Münchhof, M. *Model-Based fault detection for a hydraulic servo axis*. Doctoral thesis. TU Darmstadt, Fachbereich Elektrotechnik und Informationstechnik, Darmstadt, 2006.
5.21 Murrenhoff, H. *Servohydraulik*. Shaker Verlag, Aachen, 2002.
5.22 Roth, R. Zum Verhalten des Stellungsregelkreises. *Regelungstechnik*, 20(3):101–108, 1972.
5.23 Schaffnit, J. *Simulation und Control Prototyping zur Entwicklung von Steuergerätefunktionen für aufgeladene Nutzfahrzeug-Dieselmotoren*. Fortschr.-Ber. VDI Reihe 12, 492. VDI Verlag, Düsseldorf, 2002.
5.24 Sharif, M. and Grosvenor, R. Process plant condition monitoring and fault diagnosis. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 212(1):13–30, 1998.
5.25 Sharif, M. and Grosvenor, R. The development of novel control valve diagnostic software based on the visual basic programming language. *Proceedings of the Institution of Mechanical Engineers, Part 1: Journal of Systems and Control Engineering*, 214(2):99–127, 2000.
5.26 Song, R. and Sepehri, N. Fault detection and isolation in fluid power systems using a parametric estimation method. In *Proc. IEEE Candian Conference on Elelectrical and Computer Engineering*, volume 1, pages 144–149, Winnipeg, Manitoba, Canada, May 2002.
5.27 Stammen, C. *Condition-monitoring für intelligente hydraulische Linearantriebe*. Doctoral thesis. University of Technology, Darmstadt. RWTH Aachen, Fakultät für Maschinenwesen, Aachen, 2005.
5.28 Töpfer, S., Wolfram, A., and Isermann, R. Semi-physical modelling of nonlinear processes by means of local model approaches. In *Proc. 15th IFAC World Congress*, Barcelona, Spain, July 2002.
5.29 Watton, J. *Modelling, monitoring and diagnostic techniques for fluid power systems*. Springer, 1 edition, 2007.

Chapter 6

6.1 Dalton, T. and Patton, R. Model-based fault diagnosis of a two-pump systems. In *IFAC World Congress*, pages 79–84, San Francisco, CA, USA, 1996.

6.2 Dixon, S. *Fluid mechanics, thermodynamics of turbomachinery*. Pergamon Press, Oxford, 1966.

6.3 Fritsch, H. *Dosierpumpen*. Verlag moderne industrie, Landsberg/Lech, 1989.

6.4 Fuest, K. *Elektrische Maschinen und Antriebe*. Vieweg, Wiesbaden, 3rd edition, 1989.

6.5 Füssel, D. *Fault diagnosis with tree-structured neuro-fuzzy systems*. Fortschr.-Ber. VDI Reihe 8, 957. VDI Verlag, Düsseldorf, 2002.

6.6 Geiger, G. *Technische Fehlerdiagnose mittels Parameterschätzung und Fehlerklassifikation am Beispiel einer elektrisch angetriebenen Kreiselpumpe*. Fortschr.-Ber. VDI Reihe 8, 91. VDI Verlag, Düsseldorf, 1985.

6.7 Haus, F. *Methoden zur Störungsfrüherkennung an oszillierenden Verdrängerpumpen*. Fortschr.-Ber. VDI Reihe 8, 1109. VDI Verlag, Düsseldorf, 2006.

6.8 Hawibowo, S. *Sicherheitstechnische Abschätzung des Betriebszustandes von Pumpen zur Schadensfrüherkennung*. Doctoral thesis. Technische Universität Berlin, Berlin, 1997.

6.9 Hellmann, D. Early fault detection – an overview. *Worldpumps*, (5):2, 2002.

6.10 Hellmann, D., Kafka, D., Spath, D., and Kafka, C. Preisgünstige Überwachungssystem durch intelligente Datenanalyse. *Technische Überwachung*, 39(7/8):45–50, 1998.

6.11 Higham, E. and Perovic, S. Predictive maintenance of pumps based on signal analysis of pressure and differential pressure (flow) measurements. *Trans. of the Institute of Measurement and Control*, 23(4):226–248, 2001.

6.12 Huhn, D. *Störungsfrüherkennung an wellendichtungslosen Pumpen durch bauteilintegrierte Sensorik*. Doctoral thesis. Technische Universität, Kaiserslautern, 2001.

6.13 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2003.

6.14 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.

6.15 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.

6.16 Kafka, T., editor. *Aufbau eines Störungsfrüherkennungssystems für Pumpen der Verfahrenstechnik mit Hilfe maschinellen Lernens*. Doctoral thesis. Univ. Kaiserslautern, Kaiserslautern, 1999.

6.17 Kallweit, S. *Untersuchungen zur Erstellung wissensbasierter Fehlerdiagnosesysteme für Kreiselpumpen*. Doctoral thesis. Technische Univerisität.

6.18 Klockgether, J. and Wesser, U. *Statistische Analyse von kavitationsspezifisches Schallsignalen aus Notkühlpumpen*. Abschlussbericht des BMFT Forschungsvorhaben RS 284. Fachinformationszentrum Energie, Physik, Mathematik (FIZ), Leopoldshafen, 1981.

6.19 Kollmar, D. *Störungsfrüherkennung an Kreiselpumpen mit Verfahren des maschinellen Lernens*. Doctoral thesis. Technische Universität, Kaiserslautern, 2002.
6.20 Michaelsen, A. *Untersuchung zur automatischen Diagnose von Kreiselpumpen mit Verfahren der Signalanalyse und Mustererkennung*. Shaker, Doctoral thesis. TU Hamburg, Harburg. Aachen, 1999.
6.21 Müller-Petersen, R., Kenull, T., and Kosyna, G. Störungsfrüherkennung an Kreiselpumpen mit Hilfe der Motorstromanalyse. In *Proc. VDI-Conference Elektrisch-mechanische Antriebssysteme Innovationen-Trends-Mechatronik*, pages 441–453, VDI-Bericht Nr. 1963. Böblingen, Germany, September 2006.
6.22 Nold, S. *Wissensbasierte Fehlererkennung und Diagnose mit den Fallbeispielen Kreiselpumpe und Drehstrommotor*. Fortschr.-Ber. VDI Reihe 8, 273. VDI Verlag, Düsseldorf, 1991.
6.23 Nold, S. and Isermann, R. Model-based fault detection for centrifugal pumps and AC drives. In *11th IMEKO World Congress*, Houston, TX, USA, October 1988.
6.24 Nold, S. and Isermann, R. Die Beurteilung des Pumpenzustands durch Identifikation der Parameter von statischen und dynamischen Pumpenmodellen. In Vetter, G., editor, *Pumpen*. Vulkan-Verlag, Essen, 1992.
6.25 Nuglisch, K. *Entwicklung eines anlagenunabhängigen Störungsfrüherkennungssystems für Pumpen auf der Basis des maschinellen Lernens*. Doctoral thesis. Technische Universität Kaiserlautern, Kaiserslautern, 2006.
6.26 Pfleiderer, C. and Petermann, H. *Strömungsmaschinen*. Springer, Berlin, 7th edition, 2005.
6.27 Schlücker, E., Blanding, J., and Murray, J. Guidlines to maximize reliability and minimize risk in plants using high pressure process diaphragm pumps. In *16th Pump User Symposium*, pages 70–100, Houston, TX, USA, 1999.
6.28 Schröder, D. *Elektrische Antriebe 1*. Springer, Berlin, 1995.
6.29 VDMA Fachgemeinschaft Pumpen. *Betreiberumfrage zur Störungsfrüherkennung bei Pumpen*. VDMA, Frankfurt, 1995.
6.30 Wolfram, A. *Komponentenbasierte Fehlerdiagnose industrieller Anlagen am Beispiel frequenzumrichtergespeister Asynchronmaschinen und Kreiselpumpen*. Fortschr.-Ber. VDI Reihe 8, 967. VDI Verlag, Düsseldorf, 2002.
6.31 Wolfram, A., Füssel, D., Brune, T., and Isermann, R. Component-based multimodel approach for fault detection and diagnosis of a centrifugal pump. In *Proc. American Control Conference (ACC)*, Arlington, VA, USA, 2001.
6.32 Wolfram, A. and Isermann, R. Component-based tele-diagnosis approach to a textile machine. In *Proc. 1st IFAC Conference on Telematic Application*, Weingarten, Germany, 2001.

Chapter 7

7.1 Billmann, L. A method for leak detection and localization in gaspipelines. In *Conference on Applied Control and Identification*, Copenhagen, Denmark, 1983.

7.2 Billmann, L. *Methoden zur Lecküberwachung und Regelung von Gasfernleitungen*. Fortschr.-Ber. VDI Reihe 8, 85. VDI Verlag, Düsseldorf, 1985.

7.3 Billmann, L. and Isermann, R. Leak detection methods for pipelines. In *Proc. of the 9th IFAC Congress*, Budapest, Hungary, 1984. Pergamon Press, Oxford.

7.4 Billmann, L. and Isermann, R. Leak detection methods for pipelines. *Automatica*, 23(3):381–385, 1987.

7.5 Candy, J. and Rozsa, R. Safeguards design for a plutonium concentrator – an applied estimation approach. *Automatica*, 16(66):615–627, 1980.

7.6 Digerens, T. Real-time failure detection and identification applied to supervision of oil transport in pipelines. *Modeling, Identification and Control*, 1:39–49, 1980.

7.7 Isermann, R. Process fault detection on modeling and estimation methods - a survey. *Automatica*, 20(4):387–404, 1984.

7.8 Isermann, R. and Siebert, H. *Verfahren zur Leckerkennung und Leckortung bei Rohrleitungen Patent P2603 715.0*. 1976.

7.9 Krass, W., Kittel, A., and Uhde, A. *Pipelinetechnik*. TÜV Rheinland, Köln, 1979.

7.10 Mäncher, H., Rohrmoser, W., and Swidersky, H. *Modellbasierte Lecküberwachung von Pipelines. Report 18169. Deutsche Stiftung Umwelt*. MAGNUM Automatisierungstechnik GmbH, Darmstadt, 2002.

7.11 Siebert, H. *Evaluation of different methods for pipeline leakage monitoring (in German). PDV-Report, KfK-PDV 206*. Karlsruhe, 1981.

7.12 Siebert, H. and Isermann, R. Leckerkennung und -lokalisierung bei Pipelines durch Online-Korrelation mit einem Prozeßrechner. *Regelungstechnik*, 25(3):69–74, 1977.

7.13 Siebert, H. and Klaiber, T. Testing a method for leakage monitoring of a gasoline pipeline. *Process Automation*, pages 91–96, 1980.

Chapter 8

8.1 Freyermuth, B. Knowledge-based incipient fault diagnosis of industrial robots. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 2, pages 31–37, Baden-Baden, Germany, September 1991. Pergamon Press.

8.2 Freyermuth, B. *Wissensbasierte Fehlerdiagnose am Beispiel eines Industrieroboters*. Fortschr.-Ber. VDI Reihe 8, 315. VDI Verlag, Düsseldorf, 1993.

8.3 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.

8.4 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
8.5 Isermann, R. and Freyermuth, B. Process fault diagnosis based on process model knowledge. *Journal A, Benelux Quarterly Journal on Automatic Control*, 31(4):58–65, 1990.

Chapter 9

9.1 Altintas, Y. Prediction of cutting forces and tool breakage in milling from feed drive current measurements. *Journal of Engineering for Industry (Transactions of the ASME)*, 114(4):386–392, 1992.
9.2 Altintas, Y. *Manufacturing automation: metal cutting mechanics, machine tool vibrations, and CNC design*. Cambridge University Press, Cambridge, 2000.
9.3 Altintas, Y., Yellowley, L., and Tlusty, J. The detection of tool breakage in milling operations. *J. Engng Ind.*, 110:271–277, 1988.
9.4 Amer, W., Grosvenor, R., and Prickett, P. Machine tool condition monitoring using sweeping filter techniques. *J. Systems and Control Engineering*, 221(Part I):103–117, 2007.
9.5 Boothroyd, G. and Knight, W. *Fundamentals of machining and machine tools*. CRC Press, Boca Raton, FL, 2005.
9.6 Clark, R. A simplified instrument detection scheme. *IEEE Trans. Aerospace Electron. Systems*, 14(3):558–563, 1990.
9.7 El-Hofy, H. *Fundamentals of machining processes*. CRC Press, Boca Raton, FL, 2006.
9.8 Elbestawi, M. A., Dumitrescu, M., and Ng, E.-G. Tool condition monitoring in machining. In Wang, L. and Gao, R., editors, *Condition monitoring and control for intelligent manufacturing*, pages 55–82. Springer, London, 2006.
9.9 Ericsson, S., Grip, N., Johannson, E., Persson, L., Sjöberg, R., and Strömberg, J. Towards automatic detection of local bearing defects in rotating machines. *Mechanical Systems and Signal Processing*, 9:509–535, 2005.
9.10 Gebauer, K.-P., Maier, P., and Vossloh, M. *Statistische Fehlerursachen- und Schadensanalyse an CNC-Werkzeugmaschinen*. Institut für Produktionstechnik und spanende Werkzeugmaschinen, TU, Darmstadt, 1988.
9.11 Harris, T. *Rolling bearing analysis*. J. Wiley & Sons, New York, 4th edition, 2001.
9.12 He, X. *Modellgestützte Fehlererkennung mittels Parameterschätzung zur wissensbasierten Fehlerdiagnose an einem Vorschubantrieb*. Fortschr.-Ber. VDI Reihe 8, 354. VDI Verlag, Düsseldorf, 1993.
9.13 Isermann, R. *Identifikation dynamischer Systeme*. Springer, Berlin, 1992.
9.14 Isermann, R. Fault diagnosis of machines via parameter estimation and knowledge processing. *Automatica*, 29(4):815–835, 1993.
9.15 Isermann, R. On the applicability of model-based fault detection for technical processes. *Control Engineering Practice - CEP*, 2:439–450, 1994.

9.16 Isermann, R., editor. *Überwachung und Fehlerdiagnose - Moderne Methoden und ihre Anwendungen bei technischen Systemen*. VDI-Verlag, Düsseldorf, 1994.
9.17 Isermann, R. *Mechatronic systems – fundamentals*. Springer, London, 2nd printing edition, 2005.
9.18 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
9.19 Isermann, R., Reiss, T., and Wanke, P. Model-based fault diagnosis of machine tools. In *30th Conference on Decision and Control*, Brighton, UK, 1991.
9.20 Janik, W. *Fehlerdiagnose des Außenrund-Einstechschleifens mit Prozeß- und Signalmodellen*. Fortschr.-Ber. VDI Reihe 2, 288. VDI Verlag, Düsseldorf, 1992.
9.21 Janik, W. and Fuchs. Process- and signal-model based fault detection of the grinding process. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 2, pages 299–304, Baden-Baden, Germany, September 1991.
9.22 Jonuscheit, H., Strama, O., Henger, K., and Nass, G. Vibro-acoustics testing of combustion engines during manufacturing. *Automobiltechnische Zeitschrift – ATZ – Special Produktion*, (Nov.):46–50, 2007.
9.23 Kienzle, O. Die Bestimmung von Kräften und Leistungen an spanenden Werkzeugen und Werkzeugmaschinen. *VDI-Z*, 94(11-12):299–306, 1952.
9.24 Kolerus, J. *Zustandsüberwachung von Maschinen*. expert Verlag, Renningen-Malmsheim, 2000.
9.25 König, K., Essel, K., and Witte, L., editors. *Spezifische Schnittkraftwerte für die Zerspanung metallischer Werkstoffe*. Verein Deutscher Eisenhüttewerke, Düsseldorf, 1981.
9.26 König, W., editor. *Fertigungsverfahren Drehen, Fräsen, Bohren*. VDI, Düsseldorf, 1984.
9.27 Konrad, H. Fault detection in milling, using parameter estimation and classification methods. *Control Engineering Practice – CEP*, (4):1573–1578, 1996.
9.28 Konrad, H. *Modellbasierte Methoden zur sensorarmen Fehlerdiagnose beim Fräsen*. Fortschr.-Ber. VDI Reihe 2, 449. VDI Verlag, Düsseldorf, 1997.
9.29 Konrad, H. and Isermann, R. Diagnosis of different faults in milling using drive signals and process models. In *13th IFAC World Congress*, San Francisco, CA, USA, 1996.
9.30 Kurfess, T. R., Billington, S., and Liang, S. Y. Advanced diagnostic and prognostic techniques for rolling element bearings. In Wang, L. and Gao, R., editors, *Condition monitoring and control for intelligent manufacturing*, pages 137–165. Springer, London, 2006.
9.31 Lee, D., Hwang, I., Valente, C., Oliveira, J., and Dornfeld, D. A. Precision manufacturing process monitoring with acoustic emission. In Wang, L. and Gao, R., editors, *Condition monitoring and control for intelligent manufacturing*, pages 33–54. Springer, London, 2006.

9.32 Mikalauskas, R. and Volkovas, V. Analysis of the dynamics of a defective V-belt and diagnostic possibilites. *Proc. IMechE, Systems and Control Engineering*, 20:145–153, 2006.
9.33 Randall, R. *Frequency analysis*. Bruel & Kjaer, Naerum, 3rd edition, 1987.
9.34 Reiss, T. Model-based fault diagnosis and supervision of the drilling process. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, Baden-Baden, Germany, September 1991.
9.35 Reiss, T. *Fehlerfrüherkennung an Bearbeitungszentren mit den Meßsignalen des Vorschubantriebs*. Fortschr.-Ber. VDI Reihe 2, 286. VDI Verlag, Düsseldorf, 1992.
9.36 Sasje, E. and Mushardt, H. Instationäre Vorgänge beim Schleifen. *Industrieanzeiger*, 65:1468–1470, 1974.
9.37 Spur, G. and Stoeferle, T. *Handbuch der Fertigungstechnik*, volume 3/1. Hanser, München, 1979.
9.38 Stein, J., Colvin, D., Clever, G., and Wang, C. Evaluation of dc servo machine tool feed drives as force sensors. *J. of Dyn. Meas. and Control*, 108:279–288, 1986.
9.39 Stephenson, D. and Agapiou, J. *Metal cutting theory and practice*. CRC Press, Boca Raton, FL, 2006.
9.40 Takeyama, H. Automation developments in Japan. In *Proc. of the Third North American Metalworking Research Conference*, pages 672–685, 1975.
9.41 Tarn, J. and Tomizuka, M. Online monitoring of tool and cutting conditions in milling. *J. Engng Ind.*, 111:206–212, 1989.
9.42 Tarng, Y. and Lee, B. Use of model-based cutting simulation systems for tool breakage monitoring in milling. *Int. J. Mach. Tools Manufact.*, 32:641–649, 1992.
9.43 Tönshoff, H. and Wulfsberg, J. Developments and trends in monitoring and control of machining. *Annuals of CIRP*, 2, 1988.
9.44 Trawinski, P. and Isermann, R. Model-based fault diagnosis of a machine tool feed drive. In *21st International Symposium on Automotive Technology & Automation (ISATA)*, Wiesbaden, Germany, 1989.
9.45 Ulsoy, A. Monitoring and control of machining. In Wang, L. and Gao, R., editors, *Condition monitoring and control for intelligent manufacturing*, pages 1–32. Springer, London, 2006.
9.46 Victor, H. Schnittkraftberechnungen für das Abspanen von Metallen. *wt-Z. ind. Fertigung*, 59:317–327, 1969.
9.47 Wang, L. and Gao, R., editors. *Condition monitoring and control for intelligent manufacturing*. Springer, London, 2006.
9.48 Wang, L., Shen, W., Orban, P., and Lang, S. Remote monitoring and control in a distributed manufacturing environment. In Wang, L. and Gao, R., editors, *Condition monitoring and control for intelligent manufacturing*, pages 289–313. Springer, London, 2006.
9.49 Wanke, P. *Modellgestützte Fehlerfrüherkennung am Hauptantrieb von Bearbeitungszentren*. Fortschr.-Ber. VDI Reihe 2, 291. VDI Verlag, Düsseldorf, 1993.

9.50 Wanke, P. and Isermann, R. Modellgestützte Fehlerfrüherkennung am Hauptantrieb eines spanabhebenden Bearbeitungszentrums. *Automatisierungstechnik – at*, 40(9):349–356, 1992.
9.51 Wanke, P. and Reiss, T. Model-based fault diagnosis and supervision of the main and feed drives of a flexible milling center. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, Baden-Baden, Germany, September 1991.
9.52 Weck, M. and Brecher, C., editors. *Werkzeugmaschinen*, volume 1-5. Springer, Berlin, 2005.
9.53 Werner, G. Influence of work material on grinding forces. *Annals of the CIRP*, 27(1):243–248, 1978.
9.54 Wirth, R. Maschinendiagnose an Industriegetrieben – Grundlagen. *Antriebstechnik*, 37(10 & 11):75–80 & 77–81, 1998.
9.55 Wowk, V. *Machinery vibrations*. McGraw Hill, New York, 1991.

Chapter 10

10.1 Acklin, L. and Läubli, F. Die Berechnung des dynamischen Verhaltens von Wärmetauschern mit Hilfe von Analog-Rechengeräten. *Technische Rundschau*, 1960.
10.2 Ballé, P. Fuzzy-model-based parity equations for fault isolation. *Control Engineering Practice – CEP*, 7:261–270, 1998.
10.3 Ballé, P. *Modellbasierte Fehlererkennung für nichtlineare Prozesse mit linear-parameterveränderlichen Modellen*. Fortschr.-Ber. VDI Reihe 8, 960. VDI Verlag, Düsseldorf, 2002.
10.4 Ballé, P., Fischer, M., Füssel, D., Nelles, O., and Isermann, R. Integrated control, diagnosis and reconfiguration of a heat exchanger. *IEEE Control Systems Magazine*, 18(3):52–63, 1998.
10.5 Ballé, P. and Isermann, R. Fault detection and isolation for nonlinear processes based on local linear fuzzy models and parameter estimation. In *American Control Conference, ACC´98*, Philadelphia, PA, USA, 1998.
10.6 Goedecke, W. *Fehlererkennung an einem thermischen Prozess mit Methoden der Parameterschätzung*. Fortschr.-Ber. VDI Reihe 8, 130. VDI Verlag, Düsseldorf, 1987.
10.7 Grote, K.-H. and Feldhusen, J., editors. *DUBBEL. Taschenbuch für den Maschinenbau*. Springer, Berlin, 22nd edition, 2007.
10.8 Holman, J. *Heat transfer*. McGraw Hill, New York, 1976.
10.9 Isermann, R. Einfache mathematische Modelle für das dynamische Verhalten beheizter Rohre. *Journal Wärme*, 75:89–94, 1969.
10.10 Isermann, R. Mathematical models of steam heated heat exchangers (in German). *Regelungstechnik und Prozess-Datenverarbeitung*, 18:17–23, 1970.
10.11 Isermann, R. Einfache mathematische Modelle für das dynamische Verhalten beheizter Rohre. *Neue Technik*, (4):13–20, 1971.

10.12 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
10.13 Isermann, R. and Freyermuth, B. Process fault diagnosis based on process model knowledge. *Journal of Dynamic Systems, Measurement and Control*, 113:Part I, 620–626; Part II, 627–633, 1991.
10.14 Isermann, R. and Jantschke, H. Dynamic behavior of water- and steam-heated crossflow heat exchanger in air conditioning units (in German). *Regelungstechnik und Prozess-Datenverarbeitung*, 18:115–122, 1970.
10.15 Läubli, R. Zum Problem der Nachbildung des Verhaltens von Dampferzeugern auf Rechenmaschinen. *Technische Rundschau*, 2:35–42, 1961.
10.16 Müller-Steinhagen, H. Heat exchanger fouling. In *World Congress of Chemical Engineering*, Karsruhe/Frankfurt, Dechema, June 1991.
10.17 Neuenschwander, P. *Wärmetauscher-Ueberwachung durch Messen von Ein- und Ausgangsgrössen*. Diss. Techn. Wiss. ETH Zürich, Nr. 11576. Dietikon Juris Druck+Verlag, Zürich, 1996.
10.18 Profos, P. *Die Regelung von Dampfanlagen*. Springer, Berlin, 1962.
10.19 VDI Gesellschaft, editor. *VDI-Wärmeatlas*. VDI Verlag, Düsseldorf, 7th edition, 1994.

Chapter 11

11.1 Favre, C. Fly-by-wire for commercial aircraft: the airbus experience. *Int. Journal of Control*, 59(1):139–157, 1994.
11.2 IEC 61508. *Functional safety of electrical/electronic/programmable electronic systems*. International Electrotechnical Commission, Switzerland, 1997.
11.3 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
11.4 Lauber, R. and Göhner, P. *Prozessautomatisierung*. Springer, Berlin, 3rd edition, 1999.
11.5 Leveson, N. *Safeware. System safety and computer*. Wesley Publishing Company, Reading, MA, 1995.
11.6 Reichel, R. Modulares Rechnersystem für das Electronic Flight Control System (EFCS). In *DGLR-Jahrestagung, Deutsche Luft- und Raumfahrtkongress*, Berlin, Germany, 1999.
11.7 Reichel, R. and Boos, F. *Redundantes Rechnersystem für Fly-by-wire Steuerungen*. Bodensee-Gerätewerk, Überlingen, 1986.
11.8 Storey, N. *Safety-critical computer systems*. Addison Wesley Longman Ltd., Essex, 1996.

Chapter 12

12.1 Atkinson, G. J., Mecrow, B. C., Jack, A. G., Atkinson, D. J., Sangha, P., and Benarous, M. The design of fault tolerant machines for aerospace applica-

tions. In *Proc. IEEE International Conference on Electric Machines and Drives*, pages 1863–1869, 2005.

12.2 Beck, M. *Fault-tolerant systems - a study (in German)*. Deutsche Forschungsgesellschaft für Automatisierungstechnik und Mikroelektronik (DFAM), Frankfurt, 2008.

12.3 Beck, M., Schwung, A., Münchhof, M., and Isermann, R. Active fault tolerant control of an electro-hydraulic servo axis with a duplex-valve-system. In *IFAC Symposium on Mechatronic Systems 2010*, Cambridge, MA, USA, 2010.

12.4 Beck, M., Schwung, A., Münchhof, M., and Isermann, R. Fehlertolerante elektrohydraulische Servoachse mit Duplex-Ventilsystem. In *Automation 2010*, Baden-Baden, 2010.

12.5 Bianchi, N., Bolognani, S., and Pre, M. D. Impact of stator winding of a five-phase permanent-magnet motor on postfault operations. *IEEE Transactions on Industrial Electronics*, 55(5):1978–1987, 2008.

12.6 Bossche, D. van den. The evolution of the airbus flight control actuation systems. In *Proceedings of the 3rd International Fluid Power Conference*, Aachen, Germany, 2002.

12.7 Clark, R. State estimation schemes for instrument fault detection. In Patton, R., Frank, P., and Clark, R., editors, *Fault diagnosis in dynamic systems*, chapter 2, pages 21–45. Prentice Hall, New York, 1989.

12.8 Clarke, D. Sensor, actuator, and loop validation. *IEE Control Systems*, 15(August):39–45, 1995.

12.9 Cloyd, J. A status of the United States Air Force's more electric aircraft initiative. In *Proc. 32nd Intersociety Energy Conversion Engineering Conference IECEC-97*, volume 1, pages 681–686, 1997.

12.10 Cloyd, J. Status of the United States Air Force's more electric aircraft initiative. *IEEE Aerospace and Electronic Systems Magazine*, 13(4):17–22, 1998.

12.11 Crepin, P.-Y. *Untersuchung zur Eignung eines robusten Filterentwurfs zur Inflight-Diagnose eines elektrohydraulischen Aktuators*. PhD thesis, TU Darmstadt, Fachbereich Maschinenbau, Darmstadt, Germany, 2003.

12.12 Dilger, E. and Dieterle, W. Fehlertolerante Elektronikarchitekturen für sicherheitsgerichtete Kraftfahrzeugsysteme. *at – Automatisierungstechnik*, 50(8):375–381, 2002.

12.13 Garcia, A., Cusido, J., Rosero, J. A., Ortega, J. A., and Romeral, L. Reliable electro-mechanical actuators in aircraft. *IEEE Aerospace and Electronic Systems Magazine*, 23(8):19–25, 2008.

12.14 Goupil, P. AIRBUS state of the art and practices on FDI and FTC. In *7th IFAC International Symposium on Fault Detection, Supervision and Safety of Technical Processes. SAFEPROCESS 2009*, Barcelona, Spain, 2009.

12.15 Green, S., Atkinson, D. J., Mecrow, B. C., Jack, A. G., and Green, B. Fault tolerant, variable frequency, unity power factor converters for safety critical PM drives. *IEE Proceedings – Electric Power Applications*, 150(6):663–672, 2003.

12.16 Halbe, I. and Isermann, R. A model-based fault-tolerant sensor platform for vehicle dynamics control. In *Proceedings of the 5th Symposium in Advances*

in Automotive Control, pages 509–516, Seascape Resort Aptos, CA, USA, 2007.

12.17 Henry, M. and Clarke, D. The self-validating sensor: rationale, definitions, and examples. *Control Engineering Practice – CEP*, 1(2):585–610, 1993.

12.18 Isermann, R. Fehlertolerante mechatronische Systeme. In *VDI Tagung Mechatronik 2005*, Wiesloch, Germany, 2005.

12.19 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.

12.20 Isermann, R. and Börner, M. Characteristic velocity stability indicator for passengers cars. In *Proc. IFAC Symposium on Advances in Automotive Control*, Salerno, Italy, 2004.

12.21 Isermann, R., Lachmann, K.-H., and Matko, D. *Adaptive control systems*. Prentice Hall International UK, London, 1992.

12.22 Isermann, R., Schwarz, R., and Stölzl, S. Fault-tolerant drive-by-wire systems. *IEEE Control Systems Magazine*, (October):64–81, 2002.

12.23 Jones, R. The more electric aircraft: the past and the future? In *Proc. IEE Colloquium on Electrical Machines and Systems for the More Electric Aircraft (Ref. No. 1999/180)*, pages 1/1–1/4, 1999.

12.24 Klima, J. Analytical investigation of an induction motor drive under inverter fault mode operations. *IEE Proceedings -Electric Power Applications*, 150(3):255–262, 2003.

12.25 Krautstrunk, A. and Mutschler, P. Remedial strategy for a permanent magnet synchronous motor drive. In *8th European Conference on Power Electronics and Applications, EPE'99*, Lausanne, Switzerland, Sept 1999.

12.26 Krautstrunk, A. *Fehlertolerantes Aktorkonzept für sicherheitsrelevante Anwendungen*. Shaker Verlag, Aachen, Germany, 2005.

12.27 Kress, R. *Robuste Fehlerdiagnoseverfahren zur Wartung und Serienabhnahme elektrohydraulische Aktuatoren*. Doctoral thesis. TU Darmstadt, Fachbereich Maschinenbau, Darmstadt, 2002.

12.28 Levi, E. Multiphase electric machines for variable-speed applications. *IEEE Transactions on Industrial Electronics*, 55(5):1893–1909, 2008.

12.29 Lillo, L. de, Wheeler, P., Empringham, L., Gerada, C., and Huang, X. A power converter for fault tolerant machine development in aerospace applications. In *Proc. 13th Power Electronics and Motion Control Conference EPE-PEMC 2008*, pages 388–392, 2008.

12.30 Mesch, F. Strukturen zur Selbstüberwachung von Messsystemen. *Automatisierungstechnische Praxis – atp*, 43(8):62–67, 2001.

12.31 Moog Aircraft Group. *Redundant Electrohydrostatic Actuation System - Application: F/A-18 C/D Horizontal Stabilizer*. Moog Aircraft Group, 1996.

12.32 Moseler, O. *Mikrocontrollerbasierte Fehlererkennung für mechatronische Komponenten am Beispiel eines elektromechanischen Stellantriebs*. Fortschr.-Ber. VDI Reihe 8, 980. VDI Verlag, Düsseldorf, 2001.

12.33 Moseler, O., Heller, T., and Isermann, R. Model-based fault detection for an actuator driven by a brushless DC motor. In *14th IFAC World Congress*, volume P, pages 193–198, Beijing, China, 1999.

12.34 Moseler, O. and Isermann, R. Application of model-based fault detection to a brushless DC motor. *IEEE Trans. on Industrial Electronics*, 47(5):1015–1020, 2000.

12.35 Moseler, O. and Straky, H. Fault detection of a solenoid valve for hydraulic systems in passenger cars. In *Proceedings of the 2000 SAFEPROCESS*, Budapest, 2000.

12.36 Muenchhof, M. Condition Monitoring und Fehlermanagement für flughydraulische Servo-Achsen. In *Proceedings of the Deutscher Luft- und Raumfahrtkongress 2008*, 2008.

12.37 Muenchhof, M. Displacement sensor fault tolerance for hydraulic servo axis. In *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, 2008. International Federation of Automatic Control.

12.38 Muenchhof, M. and Clever, S. Fault-tolerant electric drives – solutions and current research activities, part I and part II. In *Proceedings of the European Control Conference 2009 – ECC 09*, Budapest, Hungary, 2009.

12.39 Müller, R., Nuber, M., and Werthschützky, R. Selbstüberwachender Durchflusssensor mit diversitärer Redundanz. *tm – Technisches Messen*, 72(4):198–204, 2005.

12.40 Münchhof, M., Beck, M., and Isermann, R. Fault-tolerant actuators and drives – structures, fault-detection principles and applications. In *7th IFAC International Symposium on Fault Detection, Supervision and Safety of Technical Processes. SAFEPROCESS 2009*, pages 1294–1305, Barcelona, Spain, 2009.

12.41 Münchhof, M. Fault management for a smart hydraulic servo axis. In *Proceedings of the Actuator 2006*, Bremen, Germany, 2006. Messe Bremen GmbH.

12.42 Navarro, R. Performance of an electro-hydrostatic actuator on the F-18 systems research aircraft. Technical Report NASA/TM-97-206224, NASA, Dryden Flight Research Center, Edwards, CA, USA, October 1997.

12.43 Oehler, R., Schoenhoff, A., and Schreiber, M. Online model-based fault detection and diagnosis for a smart aircraft actuator. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 2, pages 591–596, Hull, UK, August 1997. Pergamon Press.

12.44 Patton, R. Fault-tolerant control: the 1997 situation. In *Prepr. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, volume 2, pages 1033–1055, Hull, UK, August 1997. Pergamon Press.

12.45 Pfeufer, T. *Modellgestützte Fehlererkennung und Diagnose am Beispiel eines Fahrzeugaktors*. Fortschr.-Ber. VDI Reihe 8, 749. VDI Verlag, Düsseldorf, 1999.

12.46 Quass, S. and Schiebel, P. Aspects of future steering markets and their relevance to steering sensors. In *Proceedings of IQPC – Advanced Steering Systems*, 2007.

12.47 Raab, U. *Modellgestützte digitale Regelung und Überwachung von Kraftfahrzeugaktoren*. Fortschr.-Ber. VDI Reihe 8, 313. VDI Verlag, Düsseldorf, 1993.
12.48 Reuß, J. and Isermann, R. Umschaltstrategien eines redundaten Asynchronmotoren-Antriebssystems. In *SPS/IPC/DRIVES 2004: Elektrische Automatisierung, Systeme und Komponenten: Fachmesse & Kongress*, pages 469–477, Nürnberg, Germany, 2004.
12.49 Rosero, J. A., Ortega, J. A., Aldabas, E., and Romeral, L. Moving towards a more electric aircraft. *IEEE Aerospace and Electronic Systems Magazine*, 22(3):3–9, 2007.
12.50 Sadeghi, T. and Lyons, A. Fault tolerant EHA architectures. *IEEE Aerospace and Electronic Systems Magazine*, 7(3):32–42, 1992.
12.51 Thorsen, O. and Dalva, M. A survey of the reliability with an analysis of faults on variable frequency drives in industry. In *Proc. European Conference on Power Electronics and Applications EPE '95*, pages 1033–1038, 1995.
12.52 Weimer, J. The role of electric machines and drives in the more electric aircraft. In *Proc. IEMDC'03 Electric Machines and Drives Conference IEEE International*, volume 1, pages 11–15, 2003.
12.53 Wolfram, A. *Komponentenbasierte Fehlerdiagnose industrieller Anlagen am Beispiel frequenzumrichtergespeister Asynchronmaschinen und Kreiselpumpen*. Fortschr.-Ber. VDI Reihe 8, 967. VDI Verlag, Düsseldorf, 2002.
12.54 Wu, E. C., Hwang, J. C., and Chladek, J. T. Fault-tolerant joint development for the space shuttle remote manipulator system: analysis and experiment. 9(5):675–684, 1993.

Chapter 13

13.1 IFIP. *Proc. of the IFIP 9th World Computer Congress, Paris, France, September 19–23*. Elsevier, 1983.
13.2 Isermann, R. *Fault-diagnosis systems – An introduction from fault detection to fault tolerance*. Springer, Heidelberg, 2006.
13.3 Isermann, R. and Ballé, P. Trends in the application of model-based fault detection and diagnosis in technical processes. *Control Engineering Practice – CEP*, 5(5):638–652, 1997.
13.4 Omdahl, T., editor. *Reliability, availability and maintainability (RAM) dictionary*. ASQC Quality Press, Milwaukee, WI, USA, 1988.

Index

AC motor
  alternating current 63
  fault detection 71, 177
  fault-tolerant 294
  model 64
Actuator
  DC 306
  electrical 81, 87
  electromagnetic 81
  fault detection 85, 91
  fault diagnosis 95, 105
  fault-tolerant 301, 306
  fluidic 105
  hydraulic 105, 302
  models 88
  parameter estimation 91
  pneumatic 121
  position control 83
  quality control 89
  structure 88
  throttle valve 87
Adaptive thresholds 27
Aircraft cabin pressure valve 98
Alarms 2, 12
Assessment
  health 43
  prognostic 44
Asset management 1, 3, 5
Availability 19, 322

Change detection 26
  binary thresholds 26
Classification methods 39
Closed loop 42
Condition monitoring 1, 3
Converter 67
Cutter 245
Cutting force 243

DC motor
  brushless 98
  direct current 49
  electronic commutation 100
  model 49, 88
  test bench 49
Degradation steps 287
Dependability 21
Detection
  fault 22
Diagnostic knowledge
  analytic 208
  heuristic 208
Drilling
  Fault detection 237
  models 234
Drilling machine 234
Drive-as-sensor principle 179, 257

Electrical drives
  fault diagnosis 49
  fault-tolerant 294
  survey 50
Electrical throttle
  fault detection 91
  fault diagnosis 96
  model 88
Electromagnetic actuator

fault detection 85
model 83
position control 83
Ethylene pipeline 199

Failure 3, 17, 18
Fault 17
additive 32, 55
evaluation 16
multiplicative 32, 55
tree 61, 211
Fault detection 3, 11, 15, 22, 321
AC motor 71
actuator 85
closed loop 41
cutter 245
development 44
experimental results 55
hydraulic pump 172
knowledge-based 22
nonlinear parity equations 156
parameter estimation 34, 54, 91, 149, 156, 225
parity equations 37
process-model-based 30
SELECT 55
signal-based 24, 29, 66
state estimation 35
state observers 35
terminology 321
valve 115
vibration sensors 165
Fault diagnosis 11, 15, 22, 24, 39
actuator 105
analytical knowledge 212
classification 39
closed loop 41
cylinder 115
electrical actuator 81
electrical drives 49
error 11
fault 11, 17, 321
hierarchical 60
hydraulic pump 172
inference 41
knowledge-based 22
pumps 143
terminology 321
valve 115
Fault management 5, 6
Fault tolerance 21, 285
Fault-symptom tree 55
SELECT 55
Fault-tolerant control system 291
Fault-tolerant systems
actuator 301
basic redundant structures 285
DC actuator 306
degradation 287
drive dynamic sensors 314
dynamic redundancy 286
electrical drives 294
electrical throttle 314
examples 292
flow sensor 313
frequency converter 298
hydraulic actuators 302
sensors 308
static redundancy 286
steering angle sensor 311
Feed drives 226, 240
cascade control 228
fault detection 229
identification 229
Field-oriented control 65
Flow valves
fault detection 130, 138
Frequency converter
fault-tolerant 298
Functions 321
Fuzzy logic 41, 57, 95, 209

Grinding
models 252
Grinding machine 251

Health monitoring 5
Heat exchangers 259
fault detection dynamic behavior 275, 278
fault detection static behavior 270
models dynamic 264
models stationary 262
steam/water 274
types 259
Hydraulic servo axis 105
fault detection 115
faults 106

model 111
structure 106

Inference methods 41
Inspection 16
Integrity 21

Kalman filter 37, 317

Leak detection 181, 187
gas pipelines 195
state observers 188
Life cycle 3
Limit checking 2, 12, 25
Liquid pipelines 190
local linear model tree (LOLIMOT) 72, 163, 278

Machine tools
fault diagnosis 217
structure 217
supervision 219
Machines
drilling 234
grinding 251
milling 239
Main drive 221
Maintenance 5, 16
Malfunction 11, 17, 18, 321
Mechatronic systems 2
Methods
classification 39
fault-detection 11, 30
fault-diagnosis 11, 39
inference 41
SELECT 59
signal-analysis 29
Milling
models 239
Milling machine 239
Model 322
pneumatic valves 125
Models 221, 226, 234, 239, 252
centrifugal pump 146
cylinder 111
drilling 234
grinding 252
heat exchangers 259
milling 239
pipe systems 146
pipeline 182
robot axis 206
spool valve 111
two- and three mass 226
two-mass 221
Monitoring 1, 12
Motor
AC 63
DC 49
induction 64
multi-phase 301
MTTF (mean time to failure) 20
MTTR (mean time to repair) 20

Network models
LOLIMOT 72, 163, 278
neural 72

Output observer 35

Parameter estimation 34, 54, 223
Parity equations 37, 52
Pipe systems
models 146
Pipelines
gas 195, 202
gasoline 201
leak detection 181, 187
liquid 190
models 182
supervision 181
Plausibility
checks 28
Pneumatic actuator 121
construction 122
Pneumatic valves
fault detection 128, 130
faults 124
model 125
power electronics
fault detection 66
PRBS (pseudo-random binary signal) 55
Process 1
asset management 3
automation 1, 2
condition monitoring 1
fault management 3
management 2

models 32
product life cycles 3
supervision 1
Protection
automatic 13
systems 2
Pumps 143
centrifugal 143
diaphragm 171
drives 177
fault detection 143, 156
hydraulic 172
models 146
reciprocating 170
supervision 143
with AC motor 159
with DC motor 150

Quality control 3, 11, 89

Reconfiguration 5, 16, 316
Redundancy 21, 285
dynamic 286
static 285
Reliability 19, 322
Repair 5
Robots
faults 210
heuristic symptoms 210
industrial 205
models 206
structure 205

Safety 19, 322
Sensor
angle 311
flow 313
redundancy 308
Sensors
fault-tolerant 308
Servo axis
hydraulic 105
Signal 321
analysis 29
States 321
state estimation 36
state observer 36, 188
Structures
basic redundant 285
Supervision 1, 11, 14
basic tasks 11
data flow structure 43
Symptom 22
analytical 22, 208
generation 22
heuristic 23, 208
System
availability 19, 20
dependability 21
fault-tolerant 291
integrity 21
properties 322
reliability 19
safety 19, 20

Terminology 17, 321
fault detection 321
fault diagnosis 321
references 323
Threshold
adaptive 27
checking 12
Throttle
electrical 87, 314
Trend checking 25

Valves
faults 124
pneumatic 125
Vibration sensors 165